

STIC-Biotech/ChemLib

174981

me

From: Li, Bao-Qun
Sent: Friday, December 23, 2005 8:25 AM
To: STIC-Biotech/ChemLib
Subject: 10761006

Please do the amino acid sequence homology and interference searches against nucleotides from 155 to 835 and from 527 to 595 of SEQ ID NO: 1. Thank you.

Bao Qun Li M.D
TC 1600
Art Unit 1648
Tel. 517-272-0904
REM, 3C18
Rm. 3D24

RECEIVED
DEC 23 2005
STIC

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 1, 2005, 23:14:00 ; Search time 464.5 Seconds
(without alignments)
6082.251 Million cell updates/sec

Title: US-10-761-006A-1

Perfect score: 5907

Sequence: 1 CTCACACAACTCCACCAAG.....CCTCAGGCCAGCAGTGGA 3215

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2.1/USPTO spoel_p/US10761006/runat_01122005_113952_26862/app_query.fasta_1.3399
-DB=A Geneseq -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=ext -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10761006@cgn.1.1.746@runat_01122005_113952_26862 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_21.*

1: geneseqp1980a.*
2: geneseqp1990a.*
3: geneseqp2000a.*
4: geneseqp2001a.*
5: geneseqp2002a.*
6: geneseqp2003a.*
7: geneseqp2003bs.*
8: geneseqp2004a.*
9: geneseqp2005a.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2927	49.6	843	3 AAY54044	Aay54044 Amino aci
2	2810	47.6	540	9 ADX40768	Adx40768 HBV polym
3	2810	47.6	540	9 ADX40767	Adx40767 HBV polym
4	2807	47.5	843	9 ADX40772	Adx40772 HBV polym
5	2789	47.2	843	9 ADX40776	Adx40776 HBV polym
6	2778	47.0	845	2 ADX40777	Adx40777 HBV polym
7	2759	46.7	845	2 AAR70065	Aar70065 Hepatitis
8	2750	46.6	843	9 ADX40775	Adx40775 HBV polym
9	2737.5	46.3	539	7 ADM68212	Adm68212 GRP94 pro

10	2734.5	46.3	842	9	ADX40771	Adx40771 HBV polym
11	2688	45.5	825	9	ADX40774	Adx40774 HBV polym
12	2678.5	45.3	827	9	ADX40773	Adx40773 HBV polym
13	2675	45.3	842	4	AAE04708	Aae04708 Hepatitis
14	2675	45.3	842	4	AAG66918	Aag66918 HBV preS1
15	2654	44.9	832	9	ADX40761	Adx40761 HBV polym
16	2651	44.9	832	9	ADX40762	Adx40762 HBV polym
17	2643	44.7	843	9	ADX40760	Adx40760 HBV polym
18	2642	44.7	832	9	ADX40763	Adx40763 HBV polym
19	2628	44.5	832	9	ADX40765	Adx40765 HBV polym
20	2608	44.2	843	3	RAY44348	Ray44348 Human hep
21	2608	44.2	845	9	ADX40758	Adx40758 HBV polym
22	2607	44.1	832	9	ADX40764	Adx40764 HBV polym
23	2607	44.1	845	9	ADX40759	Adx40759 HBV polym
24	2582	43.7	843	9	ADX40766	Adx40766 HBV polym
25	1829	31.0	373	9	ADM42965	Adm42965 ADV (adef
26	1829	31.0	373	9	ADM42948	Adm42948 ADV (adef
27	1586	26.8	344	9	AEA08851	Aea08851 Hepatitis
28	1557	26.4	344	9	AEA08850	Aea08850 Hepatitis
29	1552	26.3	303	9	ADX40769	Adx40769 HBV polym
30	1552	26.3	303	9	ADX40770	Adx40770 HBV polym
31	1496	25.3	400	3	RAY54045	Ray54045 Amino aci
32	1479	25.0	309	9	AEA08832	Aea08832 Hepatitis
33	1465	24.8	281	1	AAP60163	Aap60163 Subtype a
34	1465	24.8	281	1	AAP60560	Aap60560 Hepatitis
35	1465	24.8	281	1	AAP60617	Aap60617 C-termina
36	1462	24.8	307	9	AEA08833	Aea08833 Hepatitis
37	1461	24.7	281	1	AAP70294	Aap70294 Subtype a
38	1459	24.7	400	2	AAR93801	Aar93801 Hepatitis
39	1459	24.7	400	6	ABR55863	AbR55863 HBV S-pro
40	1459	24.7	400	8	ADN49718	Adn49718 Hepatitis
41	1459	24.7	400	8	ADU74394	Adu74394 HBV S-pro
42	1457	24.7	348	7	ADG76935	Adg76935 Hepatitis
43	1456	24.6	281	2	AAR62870	Aar62870 Hepatitis
44	1456	24.6	281	4	AMM48419	AmM48419 Protein #
45	1452	24.6	389	1	AAP60794	Aap60794 Adr-cype

ALIGNMENTS

RESULT 1

AAY54044

ID AAY54044 standard; protein; 843 AA.

AC AAY54044;

DT 27-MAR-2000 (first entry)

DE Amino acid sequence of a HBV DNA polymerase protein.

KW HBV, HBV surface antigen-'S'-l33 Oon strain (Met to Thr); DNA polymerase;
KW large surface antigen; core protein; transactivating X protein;
KW hepatitis vaccine; HBV infection; hepatocellular carcinoma.

OS Hepatitis B virus.

FX Key Location/Qualifiers

FT Misc-difference 194 /note= "encoded by GAG"

FT Misc-difference 195 /note= "encoded by CAA"

FN WO9966048-A1.

PD 23-DEC-1999.

PF 19-JUN-1998; 98WO-SG000046.

PR 19-JUN-1998; 98WO-SG000046.

PA (GOVE-) GOVERNMENT REPUBLIC SINGAPORE.

PI Oon CJ, Lim GK, Zhao Y, Chen WN;

XX WPI: 2000-106104/09.
 DR N-PSDB; AAZ37088.
 XX
 PT New isolated hepatitis B virus strain, useful for, e.g. treatment of
 PT hepatitis infection.
 XX
 PS Disclosure; Page 36-39; 68pp; English.
 XX
 CC The present sequence is encoded by the genome of an isolated strain of
 CC Hepatitis B virus designated human Hepatitis B virus (HBV) surface
 CC antigen-'S'-133 Oon strain (Met to Thr). The viral genome is deposited as
 CC ECCC accession numbers P97121501, P97121502 and P97121503. The nucleotide
 CC sequence was isolated from hepatocellular carcinoma (HCC). The nucleotide
 CC sequence encodes four overlapping proteins, which are a DNA polymerase, a
 CC large surface antigen, a core protein, and a transactivating X protein.
 CC The large surface antigen differs from the wild type sequence in that it
 CC contains a Thr at position 133 of the wild type sequence instead of a
 CC Met. The proteins are used to produce antibodies. The proteins,
 CC polynucleotide and antibodies can be used for detecting the novel HBV
 CC strain. The HBV polypeptides can also be used in hepatitis vaccines. The
 CC HBV novel strain polypeptides can be used to identify compounds for
 CC treating or preventing HBV infection or hepatocellular carcinoma
 XX
 SQ Sequence 843 AA;
 Alignment Scores:
 Pred. No.: 9.31e-224 Length: 843
 Score: 2927.00 Matches: 540
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 49.55% Indels: 0
 DB: 3 Gaps: 0
 US-10-761-006A-1 (1-3215) x AAY54044 (1-843)
 QY 1 CTCACACATTCACCAAGCTCTGTAGATCCAGGCTGAGGGGCTATATTTCTCTGC 60
 DB 304 LeuHisAsnIleProProSerSerAlaArgSerGlnGlyGluGlyProIlePheSerCys 323
 QY 61 TGTGGCTCCAGTTCGGGAACAGTAAACCTGTTCGAGCTACTGCTCTCCCATATCGTC 120
 DB 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal 343
 QY 121 AATCTTCTCGAGGACTGGGACCTCGACCGAATCGGAGAACACATCAGGATTCCT 180
 DB 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
 QY 181 AGGACCCCTGCTGTGTACAGGGGGGTTTCTCGTTGACAGAGATCCTCACAATACC 240
 DB 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
 QY 241 GCAGAGTCTAGACTGTGTGAGCTTCTCTCAATTTCTAGGGGGAGCACCACGTGTTC 300
 DB 384 AlaGluSerArgLeuTrpTrpThrSerLeuAsnPhelGlyAlaProThrCysSer 403
 QY 301 TGGCCAAAATTCGAGTCCCAACCTCCATCACTACCAACCTCTGTGCTCCCAATTTG 360
 DB 404 TrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLeu 423
 QY 361 TCTGGGTATCGCTGGATGTGTCTGGGGGTTTATCATATTCCTCTCATCTCTGCT 420
 DB 424 SerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAla 443
 QY 421 ATGCTCATCTCTGTGTGTCTCTCGGACTACCAAGGTATGTGCGCGTTTGTCTCT 480
 DB 444 MetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSer 463
 QY 481 ACTTCCAGGAACATCAACCAACGACGCGGGCCATGCAAGACTCGACGACTCTCTCTCA 540
 DB 464 ThrSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSer 483
 QY 541 AGGAACCTCTACGTTTCCCTCTTGTGTCTGTACAAACCTTCGGAGCGGAACCTGCAC 600

Db 484 ArgGlyLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeu 503
 QY 601 TATTTCCCATCCCATCTCTGGGCTTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCCG 660
 DB 504 TyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPro 523
 QY 661 TTTCTCTGGCTCAGTTTACTAGTGCATTTGTTCAGTGTGTTCGTAGGGCTTTCCGCCAC 720
 DB 524 PheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHis 543
 QY 721 TGTTCGGCTTCAGTTATATATGATGATGTGGGCGCGAAGTCTGTACAAATCTT 780
 DB 544 CysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeu 563
 QY 781 GAGTCCCTTTTACCTCTATATACCAATTTTCTTTTCTTTGGGTATACATTTAAACCC 840
 DB 564 GluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPro 583
 QY 841 AATAAACCAAGCTTGGGGCTACTCCCTTAATCTCATCGGATATGTAATTTGGAAGTTGG 900
 DB 584 AsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrp 603
 QY 901 GGTACTTTTACCGCAGGAACATATTTACTAAACTCAAGCAATGTTTTCGAAACCTGCCT 960
 DB 604 GlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPro 623
 QY 961 GTAAATAGACCTATTTGATTGGAAAGTATGTCAAAGAAATTTGGGCTTTTGGCTTTGCT 1020
 DB 624 ValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAla 643
 QY 1021 GCCCTTTTACACATGTGGCTATCTCGCTTGTAGTGCCTTTATATGTCATGTATACATCT 1080
 DB 644 AlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSer 663
 QY 1081 AAGCAGGCTTTCACCTTCTCGCAACTTACAGGCTTCTGTGTAAACAAATATCTGAAC 1140
 DB 664 LysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsn 683
 QY 1141 CTTTACCCGCTTCCCGGCAACGGTCCGCTCTCTGCAAGTGTTCCTGACGCAACCCCC 1200
 DB 684 LeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPro 703
 QY 1201 ACTGGATGGGCTTGGCCATAGCCCATGCGCATGCGTGGAAACCTTTCTGGCTCTCTCTG 1260
 DB 704 ThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetAlaGlyThrPheLeuAlaProLeu 723
 QY 1261 CCGATCCATCTCGGAACCTCTAGCAGCTTGTTCGTCGAGCGGCTGAGAGCAAAA 1320
 DB 724 ProIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLys 743
 QY 1321 CTTATCGGAACCGCAACCTCTGTGTCTCTCTCGGAAATACACCTCTCTTCATCGGCTG 1380
 DB 744 LeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLeu 763
 QY 1381 CTAGGATGTGCTGCCAATCGGATCTCGCGGGAAGTCTCTTGTCTACGTCGCGTGGCG 1440
 DB 764 LeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAla 783
 QY 1441 CTGAATCCCGGAGACCCGCTCTCGGGCGGCTTTGGGGCTCTACCGTCCCTCTCTTCAT 1500
 DB 784 LeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHis 803
 QY 1501 CTGCGCTTCCCGGCCACACCGGGGGCGCACCTCTCTTTACGCGCTCTCCCGTATGTGCT 1560
 DB 804 LeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProTyrValPro 823
 QY 1561 TCTCATCTCCGGGACCGTGTGCACTTCGCTTACCTCTGCACTGCGCATGGAGACACCG 1620
 DB 824 SerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 843
 RESULT 2
 ADX40768

ID ADX40768 standard; protein; 540 AA.
 AC ADX40768;
 DT 21-APR-2005 (first entry)
 DE HBV polymerase protein #11.
 XX
 XX Immune stimulation; polymerase; enzyme.
 XX Hepatitis B virus.
 OS
 XX W02005012502-A2.
 XX
 XX 10-FEB-2005.
 XX
 XX 29-MAR-2004; 2004WO-US009510.
 XX
 XX 28-MAR-2003; 2003US-0458026P.
 XX
 XX (EPIM-) EPIMUNE INC.
 XX
 XX Baker DM, Livingston BD, Cheesnut RW, Sette A, Newman MJ;
 XX
 XX WPI; 2005-132661/14.
 DR
 XX
 XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
 PT response comprises identifying variants of a peptide epitope 8-11 amino
 PT acids in length comprising primary anchor residues of the same HLA class
 PT I binding motif.
 XX
 XX Disclosure; Page 380-385; 458pp; English.
 PS
 XX
 XX The invention relates to a method of identifying a candidate peptide
 CC epitope which induces an HLA class I CTL response against variants of the
 CC peptide epitope, comprising identifying, from a particular antigen of an
 CC infectious agent, variants of a peptide epitope comprising primary anchor
 CC residues of the same HLA class I binding motif. The method is useful for
 CC identifying a candidate peptide epitope, which induces an HLA class I CTL
 CC response against variants of the peptide epitope. This sequence
 CC represents an HBV polymerase protein used in the scope of the invention.
 XX
 XX Sequence 540 AA;
 SQ
 Alignment Scores:
 Pred. No.: 1.68e-214 Length: 540
 Score: 2810.00 Matches: 527
 Percent Similarity: 98.34% Conservative: 5
 Best Local Similarity: 97.41% Mismatches: 8
 Query Match: 47.57% Indels: 2
 DB: 9 Gaps: 0
 US-10-761-006A-1 (1-3215) x ADX40768 (1-540)
 QY 1 CTCGCAACATTCACCAAGCTCTGCTAGATCCGAGGTGAGGGCCCTATATTTTCCTGC 60
 DB 1 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGlnGlyProIlePheSerCys 20
 QY 61 TGGTGCTCCAGTCCGGAACAGTAAACCTGTCGACTACTGCTCTCCCATATCGTC 120
 DB 21 TrpTrpLeuGlnPheArgAsnSerIysProCysSerAspTyrCysLeuThrHisIleVal 40
 QY 121 AATCTTCTCGAGGACTGGGACCTCGACCGCAACATGGAGAACACAAATCAGGATTCCT 180
 DB 41 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 60
 QY 181 AGACCCCTCGTGTTCACGCGGGTTCCTGTTGACAAAGATCCTCACAATACC 240
 DB 61 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspIysAsnProHisAsnThr 80
 QY 241 GCAGACTCTAGACTCTG-GTGACTCTCTCAATTTCTAGGGGAGCACCCACCGTGTTC 299
 DB 81 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Ser 100

QY 300 CTGCCAAAATTCGCAAGTCCCAACCTCCCAATCCTCACCACCTCTTGTCTCTCAATTT 359
 DB 100 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 120
 QY 360 GTCCTGGCTATCGCTGGATGTCTGCGGGTTCCTTATCATATTCCTCTTCATCTCTGCTGC 419
 DB 120 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 140
 QY 420 TATGCCTCATCTCTTCTGTTCTTCTGCTGACTACCAAGGTATGTTGCCCTTTGCTCTC 479
 DB 140 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 160
 QY 480 TACTTCCAGGAACATCAACACCAACGACGCGGCCATCAAGACCTGCACGACTCTCTGCTC 539
 DB 160 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 180
 QY 540 AAGGAAACTCTAGCTTTCCCTCTTGTGCTGTACAAAACCTTCGACGGAACCTGCACCT 599
 DB 180 rArgAsnLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 200
 QY 600 GTATTCCCATCCCATCATCTCGGCTTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
 DB 200 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 220
 QY 660 GTTCTCTCTGGCTCAGTTTACTAGTGCATTTGTTCAGTGTTCGTAGGGCTTTCCCTCCA 719
 DB 220 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 240
 QY 720 CTGTTTGGCTTTCACTATATGATGATGATGATGATGATGATGATGATGATGATGATG 779
 DB 240 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 260
 QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTGCTTTGGGTATACATTTTAAACCC 839
 DB 260 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyLleHisLeuAsnPr 280
 QY 840 TAAATAAACCAACAGTGTGGGCTACTCCCTTAACCTTCAATGGGATATGTAATGGAGTTG 899
 DB 280 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 300
 QY 900 GGGTACTTTACCGAGGAACATATTTGACTAAACTCAAGCAATGTTTTCGAAACTGCC 959
 DB 300 pGlyThrLeuProGlnGluHisIleValGlnLysLeuLysGlnCysPheArgLysLeuPr 320
 QY 960 TGTAAATAGACTATTGATTCGMAAGTATGTCAAGAAATTTGGGTCTTTTGGGCTTTGCT 1019
 DB 320 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 340
 QY 1020 TGCCCTTTTACCAATGTGCTATCTCTGCTTGTATGCTTTTATATGATGATGATGATGAT 1079
 DB 340 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 360
 QY 1080 TAAGCAGGCTTTCACTTTCTCGCCAACTTACAAGGCTTTCTGTGTAAACAATATCTGAA 1139
 DB 360 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 380
 QY 1140 CTTTACCCCTGTCGCGCAACGGTCCGCTCTCTGCAAGTGTTCCTGACGCAACCC 1199
 DB 380 nLeuTyrProValAlaArgSerGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 400
 QY 1200 CACTCGATGGGCTTGGCCATAGCCCATGCGCATGCTGGAACCTTTCTGGCTCTCTCT 1259
 DB 400 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 420
 QY 1260 GCCGATCCATACCTCGGAACCTCTAGAGCTTGTGTTTGTCTCCAGCGGTCTGGAGCAA 1319
 DB 420 uProIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLys 440
 QY 1320 ACTTATCGGAACCGCAACTCTGTTGCTCTCTCGGAATATACACCTCTCTTCATGGGT 1379
 DB 440 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 460

QY 1380 GCTAGGCTGCTGCCACTGATCTCTGCGCGGAGCTCTTTGTCTACGTCCTCCGTCGGC 1439
 Db 460 uLeuGlyCysAlaAlaAenTrpIleLeuArgGlyThrSerPheValTrValProSerAl 480
 QY 1440 GCTGAATCCCGCGGACGACCGCTCGGGCGCTTTGGGGCTCTACCGTCCCTCTCTCA 1499
 Db 480 aLeuAenProAlaAspAspProSerArgGlyArgLeuGlyLeuTrArgProLeuLeuH1 500
 QY 1500 TCTGCGCTTCCGCGGACGACCGGCGGACCTCTCTTTACGCGGTCTCCCGTATGTGC 1559
 Db 500 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 520
 QY 1560 TTCTCATCTCCGCGGACCGTGTGACTCTGCTTCACTCTGACGTCGATGGAGACCA 1619
 Db 520 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 540
 QY 1620 G 1620
 Db 540 o 540
 RESULT 3
 ID ADX40767
 AC ADX40767; standard; protein; 540 AA.
 XX ADX40767;
 DT 21-APR-2005 (first entry)
 XX HBV polymerase protein #10.
 DE
 XX Immune stimulation; polymerase; enzyme.
 XX Hepatitis B virus.
 OS
 XX WO2005012502-A2.
 PD 10-FEB-2005.
 XX
 PF 29-MAR-2004; 2004WO-US009510.
 XX
 PR 28-MAR-2003; 2003US-0458026P.
 XX
 PA (EPIM-) EPIMUNE INC.
 XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
 PI WPI; 2005-132661/14.
 XX
 PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
 PT response comprises identifying variants of a peptide epitope 8-11 amino
 PT acids in length comprising primary anchor residues of the same HLA class
 PT I binding motif.
 XX
 PS Disclosure; Page 380-385; 458pp; English.
 XX
 CC The invention relates to a method of identifying a candidate peptide
 CC epitope which induces an HLA class I CTL response against variants of the
 CC peptide epitope, comprising identifying, from a particular antigen of an
 CC infectious agent, variants of a peptide epitope comprising primary anchor
 CC residues of the same HLA class I binding motif. The method is useful for
 CC identifying a candidate peptide epitope, which induces an HLA class I CTL
 CC response against variants of the peptide epitope. This sequence
 CC represents an HBV polymerase protein used in the scope of the invention.
 XX
 SQ Sequence 540 AA;
 Alignment Scores:
 Pred. No.: 1.68e-214 Length: 540
 Score: 2810.00 Matches: 527
 Percent Similarity: 98.34% Conservative: 5
 Best Local Similarity: 97.41% Mismatches: 8
 Query Match: 47.57% Indels: 2
 DB: 9 Gaps: 0

US-10-761-006A-1 (1-3215) x ADX40767 (1-540)
 QY 1 CTCACAAATTCACCAAGCTCTGTAGATCCAGGCTGAGGGGCTATATTTCTCTGC 60
 Db 1 LeuHisAenIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 20
 QY 61 TGTGTCTCAGTTCGGAACAGTAAACCTGTTCGACTACTGCTCTCCCATATGCTC 120
 Db 21 TrpTrpLeuGlnPheArgAenSerLysProCysSerAspTyrCysLeuThrHisIleVal 40
 QY 121 AATCTTCTCAGGACATGGGGACCTCCACGAAATGGAGAACACAACTACAGATTCT 180
 Db 41 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAenIleArgIlePro 60
 QY 181 AGGACCCCTCTCGTGTATACAGCGGGGTTTTCTCTGTGACAAAGAACTCTCAATACC 240
 Db 61 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAenProHisAenThr 80
 QY 241 GCAGATCTAGACTGTG-GTGGACTTCTCTCAATTTCTAGGGGGAGACCCACGTTCTC 299
 Db 81 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 100
 QY 300 CTGGCCAAAATTCGACAGTCCCAACCTCCAATCACCACCACTCTGTCTCTCCAATTT 359
 Db 100 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAenLe 120
 QY 360 GTCTGCTATCGCTGGATGTGTCTCGGGGTTTTATCATATTCTCTCTCTCTCTGCTGC 419
 Db 120 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 140
 QY 420 TATGCTCATCTTCTTGTGGTCTTCTGGAATACCAAGTATGTTGCCGTTTGTCTCTC 479
 Db 140 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 160
 QY 480 TACTTCCAGGACATCAACACCCAGCACGCGGGCCATGCAAGACTCTGCACACTCTCTGC 539
 Db 160 rThrSerArgAenIleAenTrpGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 180
 QY 540 AAGGAAACTCTACGTTTCCCTCTTGTGTGTGTACAAAACCTTCGACGGAACCTGCAC 599
 Db 180 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 200
 QY 600 GTATTCCTCATCCATCATCTCTGGCTTCGCAAGATTCCTATCGGAGTGGGCTCAGTCC 659
 Db 200 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 220
 QY 660 GTTTCCTCTGGCTCAGTTTACTAGTCCATTTCTTCTCAGTGGTTCTGAGGGCTTTCC 719
 Db 220 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProH 240
 QY 720 CTGTTTGGCTTTTCAGTTATATGATGATGTTGGGGCGGAAGTCTGTACAACATCT 779
 Db 240 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 260
 QY 780 TGAGTCCCTTTTACCTCTATACCAATTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTT 839
 Db 260 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAenPr 280
 QY 840 TAATAAACCAACAGTTGGGGCTACTCCCTTAACTTCTCATGGGATATGTAATTCGAAGTTG 899
 Db 280 oAenLysThrLysArgTrpGlyTyrSerLeuAenPheMetGlyTyrValIleGlySerTr 300
 QY 900 GGGTACTTTTACCAGGACACATATTCTACTAAACACTCAAGCAATGTTTTTCGAAAAC 959
 Db 300 pGlyThrLeuProGlnGluHisIleValGlnLysLeuLysGlnCysPheArgLysLeuPr 320
 QY 960 TGTAAATAGACCTATTGATTGAAAGTATGCTCAAGAAATGTGGGCTTTTGGGCTTTC 1019
 Db 320 oValAenArgProIleAenTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 340
 QY 1020 TGCCCTCTTTTACACAATGTGGCTATCTGCTCTGTGTCCTTTATATGATCATGTATACA 1079

Db 340 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 360
QY 1080 TAAGCAGGCTTTCACCTTCTCCCACTTACAGGCTTCTGTGTAAACATACTCTGAA 1139
Db 360 rlySGlnAlaPheThrPheSerProThrTyrIlySaIaPheLeuCySLySGlnTyrLeuAs 380
QY 1140 CTTTACCCCGTTCGCCGCAACGGTCCGGTCTCTCGCAAGTGTGTGTGACGCAACCCC 1199
Db 380 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaPheAlaThrPr 400
QY 1200 CACTGGATGGGCTTGGCCATAGGCGCATAGCGCATGGCTGGAACCTTCTGTGGCTCTCT 1259
Db 400 oThrGlyTyrGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe 420
QY 1260 GCCGATCATACTCGGGAACCTCTAGCAGCTTGTCTCGCAGCCGCTCTGGAGCAA 1319
Db 420 uProIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy 440
QY 1320 ACTTATCGGAACCGCAACTCTGTGTCTCTCTCGGAAATACACTCTCTTCCATGGCT 1379
Db 440 sleulleGlyThrAspAsnSerValValLeuSerArgIlyThrSerPheProTrpLe 460
QY 1380 GCTAGGGTGTGTCGCAACTGGATCTCGCGGGACGTCCTTGTCTACGTCCCGTCGGC 1439
Db 460 uLeuGlyCySaIaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 480
QY 1440 GCTGAATCCCGGAGCAGCCCTCTCGGGCCGTTGGGGCTCTACCGTCCCTCTCTTCA 1499
Db 480 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 500
QY 1500 TCTGCGCTTCCGGCGGACCGCGGCGGACCTCTCTTTACGGGCTCCCGGTATGTGCC 1559
Db 500 sleuProPheArgProThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 520
QY 1560 TTCTCATCTCGCGGACCGTGTGCACTTCGCTTCACTCTGCACTGCGATGAGACACC 1619
Db 520 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 540
QY 1620 G 1620
Db 540 o 540
RESULT 4
ID ADX40772 standard; protein; 843 AA.
XX AC ADX40772;
XX DT 21-APR-2005 (first entry)
XX DE HBV polymerase protein #15.
XX KW Immune stimulation; polymerase; enzyme.
XX OS Hepatitis B virus.
XX PN WO2005012502-A2.
XX PD 10-FEB-2005.
XX PF 29-MAR-2004; 2004WO-US009510.
XX PR 28-MAR-2003; 2003US-0458026P.
XX PA (EPIM-) EPIMUNE INC.
XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.
PT Identifying a candidate peptide epitope, which induces a HLA class I CTL response comprises identifying variants of a peptide epitope 8-11 amino acids in length comprising primary anchor residues of the same HLA class

I binding motif.
Disclosure; Page 380-385; 458pp; English.
The invention relates to a method of identifying a candidate peptide epitope which induces an HLA class I CTL response against variants of an infectious agent, comprising identifying, from a particular antigen of an infectious agent, variants of a peptide epitope comprising primary anchor residues of the same HLA class I binding motif. The method is useful for identifying a candidate peptide epitope, which induces an HLA class I CTL response against variants of the peptide epitope. This sequence represents an HBV polymerase protein used in the scope of the invention.
SQ Sequence 843 AA;
Alignment Scores:
Pred. No.: 3,31e-214 Length: 843
Score: 2807.00 Matches: 528
Percent Similarity: 98.33% Conservative: 3
Best Local Similarity: 97.78% Mismatches: 8
Query Match: 47.52% Indels: 2
DB: 9 Gaps: 0
US-10-761-006A-1 (1-3215) x ADX40772 (1-843)
QY 4 CACACACATTCACCAAGCTCTGTAGATCCAGGGTGAGGGCCCTATATTTTCTGTCTGG 63
Db 305 HisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCysTrp 324
QY 64 TGGCTCCAGTTCGCGAACAGTAAACCCCTGTTCCGACTACTGCCTCTCCCATATCGTCAAT 123
Db 325 TrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAsn 344
QY 124 CTTCTCGAGGACTGGGACCTCGACCGAACATCGAGAACACATCAGGATTCCTAGG 183
Db 345 LeuLeuGluAspTyrGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArg 364
QY 184 ACCCTCTCTGTGTACAGGGGGGTTTCTCGTGTGACAGAAATCCTCACAATACCGCA 243
Db 365 ThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrThr 384
QY 244 GAGTCTAGACTCTG-GTGGACTTCTCAATTTTCTAGGGGAGCACCCAGCTGTCTCTG 302
Db 385 GluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-SerTr 404
QY 303 GCCAAATTCGAGTCCCAACCTCCAAATCACTCACCACCTCTGTGTCTCCAAATTTGTC 362
Db 404 pProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSe 424
QY 363 CTGGCTATCGTGGATGTCTCGGGGGTGTATATCATATTCCTCTCATCTGTCTGTCTAT 422
Db 424 rTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisPProAlaAlaMe 444
QY 423 GCCTCATCTCTGTGTGTCTTCTGACTACCAAGGTATGTGTCCCTTTGTCTCTCTAC 482
Db 444 tProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerTh 464
QY 483 TTCAGGGAACATCAACCAACAGCAGCGGGGCCATGCAAGACCTGCAAGCTCTCTGCTCAAG 542
Db 464 rSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSerAr 484
QY 543 GAAACTCTACGTTTCCCTCTGTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 602
Db 484 gAsnLeuTyrValSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 504
QY 603 TTCCCATCCCATCATCTCTGGGCTTTCGCAAGATTCCTATGGAGTGGGCGCTCAGTCCGTT 662
Db 504 rSerHisProIleIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPh 524
QY 663 TCTCCTGCTCAGTTTACTAGTGCATTTGTTCAGTGGTTCGTAGGGTTCCTCCCACTG 722
Db 524 eLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHisCy 544

QY 723 TTTGGCTTTCAGTTATATGATGATGCTGTTATTTGGGGCGGAAGTCTGTACAAATCTTGA 782
Db 544 sLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLeuGI 564
QY 783 GTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCTTAA 842
Db 564 uSerLeuPheThrSerIleThrAsnPheLeuSerLeuGlyIleHisLeuAsnProAs 584
QY 843 TAAAAACCAACGTTGGGGCTACTCCCTTAATCTTCATGGGATATGTAATTTGGAAGTTGGGG 902
Db 584 nLysThrLysArgTyrGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrpGI 604
QY 903 TACTTTACCCGAGAACATATTGTACTTAAACCTCAAGCAATGTTTTCGAAAACTCGCTGT 962
Db 604 yThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuProVa 624
QY 963 AAATAGACCTATTGATTGGAAAGTATGTCAAGAAATTTGGGCTTTTGGGCTTTCTGCTGC 1022
Db 624 lAsnSerProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAl 644
QY 1023 CCCTTTTACAAATGTGGCTATCTCTGCTTGATGCTCTTTATATGCATGTATACAATCTAA 1082
Db 644 aProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerLy 664
QY 1083 GCAGGCTTTCACCTTCTCGCAACTTACAAGGCTTTCTGTGTAAACAATATCTGAACCT 1142
Db 664 sGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsnLe 684
QY 1143 TTACCCCGTTGGCCGCAACGGTCCGCTCTCTCCAAAGTCTTCTGACGCAACCCCCAC 1202
Db 684 uTyrProValAlaArgGlnArgSerGlyLeuGlnValPheAlaAspAlaThrProTh 704
QY 1203 TGGATGGGCTTGGCCATAGGCCATCAGCATGGCTGGAACTTTCTGCTCTCTCTGCC 1262
Db 704 xGlyTyrGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLeuPr 724
QY 1263 GATCCATCTGCGGAACCTCTAGCAGTGTGTTTGTCTCGAGCGGTCTGGAGCAAACT 1322
Db 724 oIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLysLe 744
QY 1323 TATCGGAACGACAACTCTGTGTCTCTCTCGAAATACACCTCTTTCATGGCTGCT 1382
Db 744 uIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLeuLe 764
QY 1383 AGSGTGTCTGCCAACTGGATCTCTGCGCGGACGTCTCTTTGTCTACGTCCCGTCGCGCT 1442
Db 764 uGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAlaLe 784
QY 1443 GAATCCCGGAGCAGCCGCTCTCGGGCCGTTTGGGGCTCTACCGTCCCTTCTTCATCT 1502
Db 784 uAsnProAlaAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuLeuLe 804
QY 1503 GCCGTTCCGCGCCACCAACGCGGCGCACTCTCTTTACGCGGTCTCCCGPATGTGCTTC 1562
Db 804 uProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValProSe 824
QY 1563 TCATCTGCCGACCGTGTGCACTTCTGCTTCCACTCTGACGTGCGATGGAGACCCG 1620
Db 824 xHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 843
RESULT 5
ADX40776 standard; protein; 843 AA.
XX ID
AC ADX40776;
XX
XX
DT 21-APR-2005 (first entry)
XX
DE HBV polymerase protein #19.
XX
KW Immune stimulation; polymerase; enzyme.
XX
QS Hepatitis B virus.

XX
PN W02005012502-A2.
XX
PD 10-FEB-2005.
XX
XX 29-MAR-2004; 2004WO-US009510.
XX
XX 28-MAR-2003; 2003US-0458026P.
XX
XX (EPIM-) EPIMMUNE INC.
XX
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
PI WPI; 2005-132661/14.
XX
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX
XX Disclosure; Page 380-385; 458pp; English.
XX
XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HBV polymerase protein used in the scope of the invention.
XX
XX Sequence 843 AA;
SQ
Alignment Scores:
Pred. No.: 8,98e-213 Length: 843
Score: 2789.00 Matches: 523
Percent Similarity: 98.15% Conservative: 8
Best Local Similarity: 96.67% Mismatches: 9
Query Match: 47.22% Indels: 2
DB: 9 Gaps: 0
US-10-761-006A-1 (1-3215) x ADX40776 (1-843)
QY 1 CTCACAAATTCACCAAGCTCTGTAGATCCAGGGTGAAGGGCTATATTTCTCTGC 60
Db 304 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGlnGlyProIlePheSerCys 323
QY 61 TGTGTGCTCCAGTTCGCGAACAGTAAACCTGTTCGACTACTGCTCTCCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
QY 121 AATCTTCTCGAGACTGCGGACCTCGACCGAACATGGAGAACACAACTCAGGATTCCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGACCCCTGCTGCTGTATACAGCGGGGTTTTCTCGTTGACAAGAAATCTCAATATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 GCAGACTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGGAGACCCACGCTGTTTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
QY 300 CTGGCCAAAATTCGCAAGTCCCAACCTCCAATCAGTCCCAACCTCTGTCTCTCAATTTT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
QY 360 GTCTGCTGCTATCGCTGATGTGTCTCGGGGTTTTATCATATTTCTTCTTCATCTCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaPheTyrHisIleProLeuHisProAlaAl 443
QY 420 TATGCTCATCTTCTTGTGTGTTCTTCTGACTTACCAAGGTATGTTGCCGTTTGTCTTC 479

```
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
Qy 480 TACTTCAGGAACATCAACACAGCAGCGGGCCATGCAAGACCTGCAGACCTCTCTCTC 539
Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
Qy 540 AAGGAAACTCTACCTTTCCCTCTCTGTTGCTGTACAAAACCTTCGAGCGGAACTGCACATT 599
Db 483 rArgAsnLeuTyrValSerLeuLeuValTyrLysThrPheGlyArgLysLeuHisLe 503
Qy 600 GTATTCCCATCCCATCTCTGGGCTTTTCGCAAGATTCTATGGAGTGGGCTTCAGTCC 659
Db 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Qy 660 GTTCTCTCTGCTCAGTTTACTAGTGCATTTGTTTCAGTGGTTCTGTAGGGCTTTCCCCCA 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 543
Qy 720 CTGTTTGGCTTTCAGTTATATGGATGATGCTGTTATGGGGCGAAGTCTGTACAACTCT 779
Db 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
Qy 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
Db 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
Qy 840 TAATAAAACCAAAAGCTTGGGCTACTCCCTTAACCTTCATGGGATATGTAATGGAAGTTG 899
Db 583 oAsnLysThrLysArgTyrGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
Qy 900 GGGTACTTACCGCAGGACATATGTACTAACTCAAGCAATGTTTTCGAAACCTGCC 959
Db 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
Qy 960 TGTAAATAGACTATTGATGGAAAGTATGTCAAGAGATTTGGTCTTTTGGGCTTTGC 1019
Db 623 oValAsnArgProIleAspTyrLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
Qy 1020 TGCCCTTTTACAAATGTGCTATCTCGCTTGTATGCTTGTATGATGCTTTATGATGATCAATC 1079
Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
Qy 1080 TAAGCAGGCTTTCATTTCTCGCAACTTACAAAGGCTTTCTGTGTAAACAATATCTCAA 1139
Db 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuLysGlnGlnTyrLeuHi 683
Qy 1140 CTTTACCCCGTTCGCGCAACGCTCGGCTCTCTGCAAGTGTGTGTGACGCAACCCC 1199
Db 683 sLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheGlyAspAlaThrPr 703
Qy 1200 CACTGGATGGGCTTGGCCATAGGCCATCAGCGCATGCTGGAACCTTTCTGGCTCTCT 1259
Db 703 oThrGlyTyrGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 723
Qy 1260 GCCGATCATACTCGGGAACCTCTAGCAGCTGTGTTGCTCGCAGCGCTCTGGAGCAA 1319
Db 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgGlyAlaLys 743
Qy 1320 ACTTATCGGAACCGCAACTCTGTTGCTCTCTCGGAATACACCTCTCTTCCATGCT 1379
Db 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTyrLe 763
Qy 1380 GCTAGGGTGTGTCGCAACTGGATCTCTCGCGGAGCGTCTTTTGTCTAGCTCCCTCGGC 1439
Db 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
Qy 1440 GCTGAATCCCGGAGCAGCCCTCTCGGGCGGTTTGGGCTCTACCGTCCCTCTTCA 1499
Db 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
Qy 1500 TCTGCGGTTCCGCGCAGCAGCGGCGCACCTCTCTTACGCGGCTCTCCCGCTATGTC 1559
Db 803 sLeuProPheGlnProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
```

```
Qy 1560 TTCTCATCTCCGACCGTGTGCACTTCGCTTCACTCTGCTCTGACGTCGATGGAGCACACC 1619
Db 823 oSerHisLeuProValArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
Qy 1620 G 1620
Db 843 o 843

RESULT 6
ADX40777
ID ADX40777 standard; protein; 845 AA.
XX
AC ADX40777;
XX
DT 21-APR-2005 (first entry)
XX
DE HBV polymerase protein #20.
XX
KW Immune stimulation; polymerase; enzyme.
XX
OS Hepatitis B virus.
XX
PN WO2005012502-A2.
XX
PD 10-FEB-2005.
XX
PF 29-MAR-2004; 2004WO-US009510.
XX
PR 28-MAR-2003; 2003US-0458026P.
XX
PA (EPIM-) EPIMUNE INC.
XX
PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX
DR WPI; 2005-132661/14.
XX
PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX
PS Disclosure; Page 380-385; 458pp; English.
XX
CC The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HBV polymerase protein used in the scope of the invention.
XX
SQ Sequence 845 AA;
```

```
Alignment Scores:
Pred. No.: 6.74e-212 Length: 845
Score: 2778.00 Matches: 521
Percent Similarity: 97.41% Conservative: 6
Best Local Similarity: 96.30% Mismatches: 13
Query Match: 47.03% Indels: 2
DB: 9 Gaps: 0
```

```
US-10-761-006A-1 (1-3215) x ADX40777 (1-845)
Qy 1 CTCACAAACATTCCACCAAGCTCTGTAGATCCCAAGGTGAGGGCCCTATATTTCTCTGC 60
Db 306 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGluGlyPro***PheSerCys 325
Qy 61 TGGTGGCTCCAGTTCGGAACAGTAAACCTGTTCCGACTACTGCTCTCCCATATCGTC 120
Db 326 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal 345
```


CC used, particularly ex vivo, to stimulate CTL cells. These cells can be
 CC reintroduced into patients who have chronic or acute HBV infections or
 CC are carriers, especially in treatments to prevent conversion from acute
 CC to chronic infections. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 845 AA;

Alignment Scores: 2.19e-210 Length: 845
 Pred. NO.: 2759.00 Matches: 516
 Score: 97.23% Conservative: 10
 Best Local Similarity: 95.38% Mismatches: 14
 Query Match: 46.71% Indels: 2
 DB: Gaps: 0

US-10-761-006A-1 (1-3215) x AAR70065 (1-845)

Qy 1 CTCACAACTTCCACCAAGCTCTGCTAGATCCAGGGGTGAGGGCCCTATATTTTCCTGC 60
 Db 306 LeuHisAnlileProProAsnSerAlaArgSerGlnSerGluGlyProValPheSerCys 325
 Qy 61 TGGTGGCTCCAGTTCGGGAAACAGTAACCCCTGTTCCGACTACTCGCTCTCCCATATGTC 120
 Db 326 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 345
 Qy 121 AATCTTCTCGAGGACTGGGACCCCTGCACCGAAACATGAGAGAACACATCAGGATTCCCT 180
 Db 346 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 365
 Qy 181 AGGACCCCTGCTCGTGTTCACAGGGGGGTTTTCTCTGTTGACAGAAATCCTCACAATACC 240
 Db 366 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 385
 Qy 241 GCAGAGTCTAGACTCTG-GTGACTTCTCAATTTCTAGGGGAGACCCACGTTGTC 299
 Db 386 ThrGluSerArgLeuValAspPheSerGlnPheSerArgGlySerThrArgVal-Se 405
 Qy 300 CTGGCCAAATTCGAGTCCCAACCTCAATCACTACCAACCTCTGCTCTCCCAATTT 359
 Db 405 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 425
 Qy 360 GTCTGGCTATCTGCTGGATGTCTGCGCGTTTTTATCATATTCCTCTTCATCTCTGCTGC 419
 Db 425 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 445
 Qy 420 TATGCTCATCTCTCTGTGTGTTCTCTGACTACCAAGGTATGTTGCCGTTGTCTCTC 479
 Db 445 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 465
 Qy 480 TACTTCGAGGAACATCAACACACGACGCGGCGCATGCAAGACCTGCGAGCTCTGCTC 539
 Db 465 rAsnSerArgIleIleAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 485
 Qy 540 AAGGAACTCTAGCTTTCCCTCTCTGCTGTGACAAAACCTTCGAGCGGAACTGCACCT 599
 Db 485 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 505
 Qy 600 GTATTCCCATCCCATCATCTGGGCTTTTCGCAAGATTCTTATGGAGTGGGCTCAGTCC 659
 Db 505 uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 525
 Qy 660 GTTCTCTCTGGCTCAGTTTACTAGTGCACATTTGTTGTCAGTGGTTCGTAGGGCTTCCCA 719
 Db 525 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 545
 Qy 720 CTGTTTGGCTTTCAGTTATATGATGATGTGTTATTTGGGGGAGATCTGTACACATCT 779
 Db 545 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 565
 Qy 780 TCAGTCCCTTTTATCTCTTATACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
 Db 565 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuLeuSerLeuGlyIleHisLeuAsnPr 585

Qy 840 TAATAAAACCAACAGTGGGGCTACTCCCTTAACATTCATCGGATATGTAATGGAAGTTG 899
 Db 585 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 605
 Qy 900 GGGTACTTTACCGGAGGACATATTGTACTAAACATCAAGCAATGTTTTCGAAACCTGCC 959
 Db 605 pGlyThrIleProGlnGluHisIleValGlnLysIleLysGlnCysPheArgLysLeuPr 625
 Qy 960 TGTAAATAGACCTATTGATTGGAAAGTATGTCACAAAGAAATGTGGGTCTTTTGGCTTTC 1019
 Db 625 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 645
 Qy 1020 TGCCCTTTTACACAATGTGCTATCTCTGCTGATCCCTTATATATCATGTATACATC 1079
 Db 645 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 665
 Qy 1080 TAAGCAGGCTTTCACCTTCTCGCCAACTTACAGGCGCTTTCTGTGTAACAATATCTGAA 1139
 Db 665 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 685
 Qy 1140 CTTTACCCCGTTGCGCGCAACGGTCCGGTCTCTGCAAGTGTTCGTGACGCAACCCC 1199
 Db 685 nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 705
 Qy 1200 CACTGGATGGGGCTTGCCCATAGCCATCAGCGCATCGCTGGAACCTTTCTGGCTCTCT 1259
 Db 705 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 725
 Qy 1260 GCCATCCATCTCGCGAATCTCTAGCAGCTTGTTCGTCGCGAGCGGTCTGGAGCAA 1319
 Db 725 uProIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLys 745
 Qy 1320 ACTTATCGGAACCGCAACTCTGTGTCTCTCGGAAATACACCTCTCTTTCATGCT 1379
 Db 745 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 765
 Qy 1380 GCTAGGTGTGCTCCCAACTCGATCTCGCGGAGCGTCTCTTCTGTACGTCGCGTCCGC 1439
 Db 765 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 785
 Qy 1440 GCTGAATCCCGCGAGCACCGCTCTCGGGCGCGTTTGGGGCTCTACCGTCCCTCTTCA 1499
 Db 785 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuAr 805
 Qy 1500 TCTGCGTTCGCGCGGACCGCGGCGCACCTCTCTTACGCGGTCTCCCGTATGTGCC 1559
 Db 805 gLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 825
 Qy 1560 TTCTCATCTGCGGACCGTGTGCACTTCTGCTTTCACCTCTGCACGTCGCGATGGAGACC 1619
 Db 825 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTyrArgProPr 845
 Qy 1620 G 1620
 Db 845 o 845
 RESULT 8
 ID ADX40775 standard; protein; 843 AA.
 AC ADX40775;
 XX 21-APR-2005 (first entry)
 DT HBV polymerase protein #18.
 DE Immune stimulation; polymerase; enzyme.
 KW Hepatitis B virus.
 OS WO2005012502-A2.
 PN 10-FEB-2005.
 PD

XX 29-MAR-2004; 2004WO-US009510.
PF
XX 28-MAR-2003; 2003US-0458026P.
PR
XX (BPIIM-) EPIMMUNE INC.
PA
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
PI WPI; 2005-132661/14.
DR
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX
XX Disclosure; Page 380-385; 458pp; English.
XX
XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HBV polymerase protein used in the scope of the invention.
XX
XX Sequence 843 AA;
SQ

Alignment Scores:
Pred. No.: 1.14e-209 Length: 843
Score: 2750.00 Matches: 517
Percent Similarity: 97.04% Conservative: 8
Best Local Similarity: 95.56% Mismatches: 15
Query Match: 46.55% Indels: 2
DB: 9 Gaps: 0

US-10-761-006a-1 (1-3215) x ADX40775 (1-843)

QY 1 CTCACAAATTCACCAAGCTCTGTAGATCCAGGGTGAGGGGCTATATTTTCTCTGC 60
DB
304 LeuHisAsnIleProSerSerAlaArgSerLysSerGluGlyProLeuPheProCys 323

QY 61 TGTGGCTCAGTTCGGGAACAGTAACCTGTTCCGACTACTGCTCTCCCATATCGTC 120
DB
324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343

QY 121 AATCTTCTCAGAGACTGGGGACCTGCACCGAACATCGGAGAACACACATCAGGATTCCT 180
DB
344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyHisAsnIleArgIlePro 363

QY 181 AGGACCCCTCTCGTGTATACAGCGGGGTTTTCTCGTTGACAGAATCCCTCAATAACC 240
DB
364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383

QY 241 GCAGAGTCTAGACTCTG-GRGACTTCTCTCAATTTCTAGGGGGAGACCCACGTTGTC 299
DB
384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnThrGlnVal-Se 403

QY 300 CTGGCCAAATTCGAGTCCCAACCTCCCAATCACCACCACTCTGTCTCTCCCAATTT 359
DB
403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 423

QY 360 GTCTCGCTATCGCTGATGTCTCGGGGTTTTATCATATTTCTCTTCATCTCTGCTGC 419
DB
423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaH 443

QY 420 TATGCTCATCTCTTGTGTTCTTCTGACTACCAAGGTATGTTGCCGTTTGTCTTC 479
DB
443 aMetProHisLeuLeuValGlySerSerGlyLeuGlnArgTyrValAlaArgLeuSerSe 463

QY 480 TACTTTCAGGAACATCAACACACGACCGGGCCATGCAAGACCTCAGACCTCTCTGCTC 539
DB

DB 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
QY 540 AAGAAACTTACGTTTCCCTCTGTGTGTGTACAAAACTTCGGAGGAAAACTGCACATT 599
DB
483 rLysHisLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
QY 600 GTATTCCCATCCCATCATCTGGGCTTCGCAAGATTCCTATGGAAGTGGGCTCAGTCC 659
DB
503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
QY 660 GTTTCCTCTGGGCTCAGTTTACTAGTGCCTTTTGTTCAGTGGTTCGTAGGCTTCCCCCA 719
DB
523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValCysArgAlaPheProH 543
QY 720 CTGTTTGGCTTTCAGTTATATGATGATGTGTATTGGGGGGAAGTCTGTACAAATCT 779
DB
543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAGTCCCTTTTACTCTCTATTACCAATTTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
DB
563 uGluSerLeuPheThrSerIleThrAsnPheMetLeuSerLeuGlyIleHisLeuAsnPr 583
QY 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACITTCATGGGATATGTAATTGGAAAGTTG 899
DB
583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
QY 900 GGGTACTTTACCGCAGGAACATATTGTACTTAAACCTCAAGCAATGTTTTCGAAAACTGCC 959
DB
603 pGlyThrLeuProGlnGluHisIleValGlnLysLeuLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAAATACACCTATTGATTGGAAAGTATGTCAAAGAAATGTTGGGCTTTTGGGCTTTC 1019
DB
623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGCCCTTTTACACAATGCGCTATCTCTGCTTGATGCTCTTTATATGATGATATACAATC 1079
DB
643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
QY 1080 TAAGCAGGCTTTTCACTTTCTCGCCAACTTCAAGGCTTTTGTGTGTAAACAATATCTGAA 1139
DB
663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 683
QY 1140 CCTTTACCCGGTTCGCCGGAACGGTCCGCTCTCTGCGAAGTGTTCGTGAGCGCAACCCC 1199
DB
683 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
QY 1200 CACTGGATGGGCTTGGCCATAGGCCATCAGCGCATGCGCTGGAACCTTTCTGCTCTCTCT 1259
DB
703 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 723
QY 1260 GCGATFCCATCTGCGGAACCTCTAGCAGCTTGTGTTTGTCTGCGAGCGGCTCTGGAGCAA 1319
DB
723 uProIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaTh 743
QY 1320 ACTTATCGGAACCGCAACCTCTGTCTCTCTCGGAATATACACCTCTTTCATGGCT 1379
DB
743 rLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
QY 1380 GTTAGGGTGTCTGCGCAACTGGATCTCTGCGGAGCGTCTTGTGTCTAGCTCCCGCTCGG 1439
DB
763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
QY 1440 GCTGAATCCCGGAGCGACCCCTCTCGGGCGCTTGGGGCTCTACCGTCCCTTCTTCA 1499
DB
783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuAr 803
QY 1500 TCTGCGGTTCCGGCGCACCCGCGGCGCTCTCTTTACGGGCTCTCCCGCTATGTGCC 1559
DB
803 gLeuSerPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
QY 1560 TTCTCATCTCGCGGACCGGTGTGCATTTCCCTTTCCACCTCTGCGAGCTGCGATGAGAGCACCC 1619
DB
823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843

QY	1620 G 1620	61	TGCTGGCTCCAGTTCGGAAACAGTAAACCTGTTCGACTACTGCTCTCCCATATGCTC	120
DB	843 o 843	21	TriPLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal	40
RESULT 9		121	AATCTTCTCGAGACTGGGGACCTCCACCGAATCGGAGAACACACATCAGGATTCCT	180
ID	ADM68212	41	AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleAArgIlePro	60
XX	ADM68212; standard; protein; 539 AA.	181	AGGACCCCTGCTGCTGTACAGCGGGGGTTCCTGCTTGCATGACAGAAATCCTCAATPACC	240
XX	ADM68212;	61	ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr	80
DT	03-JUN-2004 (first entry)	241	GCAGAGCTAGACTCTG-GRGACTTCTCTCAATTTTCTAGGGGGAGCACCCAGCTGTC	299
XX	GRP94 protein related Hepatitis B virus protein, SEQ ID NO 4.	81	ThrGluSerArgLeuValAlaAspPheSerGlnPheSerArgGlySerThrHisVal-Se	100
XX	GRP94; chaperone; DNA polymerase; Hepatitis B virus; HBV.	300	CTGGCCAAATTCGAGTCCCAACCTCCAACTCACCAACCTCTGTGCTCCCAATTT	359
XX	Hepatitis B virus.	100	rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe	120
PN	KR2002061721-A.	360	GTCTGGCTATCGCTGGATGTCTGGCGGCTTTTATCATATTCCTCTTCATCTGCTGC	419
XX	25-JUL-2002.	120	uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl	140
XX	17-JAN-2001; 2001KR-00002647.	420	TATGCTCATCTTCTGCTGCTTCTGACTACCAAGTATGCTGCTGCTGCTGCTC	479
PF	17-JAN-2001; 2001KR-00002647.	140	aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe	160
XX	(MEDE-) MEDEXBIO CO.	480	TACTTCCAGGAACATCAACACGAGCAGCGGGCCATCGACACCTGACACCTCTGCTC	539
XX	Cho YH, Kim SS, Noh HM, Shin HJ;	160	rThrSerArgAsnIleAsnHisGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe	180
DR	WPI; 2003-117827/11.	540	AAGGAAACTCTAGCTTTCCTCTGTTGCTGTACAAAACCTTCGAGCGGAACTGCATTT	599
DR	N-PSDB; ADM68211.	180	rArgAsnLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe	200
XX	Grp94 acting as chaperone of DNA polymerase of hepatitis b virus and	600	GTATTCCTCCATCCATCCTGGGCTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC	659
PT	expression vector system for preparing a DNA polymerase in active state	200	uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyGlyGlyLeuSerPr	220
PT	of hepatitis b virus.	660	GTTTCTCTGCTGCTAGTTACTAGTGCATTTGTTGCTGCTGCTGCTGCTGCTGCTGCT	719
XX	Disclosure; SEQ ID NO 4; 15pp; Korean.	220	oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi	240
CC	The invention relates to a novel GRP94 protein. The GRP94 protein acts as	720	CTGTTTCTGCTTTCAGTTATATGATGATGATGATGATGATGATGATGATGATGATGAT	779
CC	a chaperone of DNA polymerase of hepatitis B virus. The invention further	240	scysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe	260
CC	relates to: an expression vector system for preparing a DNA polymerase in	780	TGAGTCCCTTTTACCTCTATTACCAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT	839
CC	the active state of a hepatitis B virus, a transformant and a	260	uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr	280
CC	manufacturing method of DNA polymerase of active hepatitis B virus using	840	TAATAAACCAACGTTGGGCTACTCCCTTAACCTCATGGGATATGTAATGGAATGG	899
CC	the transformant, and rapidly screening potential matter for the	280	oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr	300
CC	prevention and treatment of diseases associated with hepatitis B virus.	900	GGGTACTTTTACCGCAGGAACATATTGTACTAAAACCTCAAGCAATGTTTTTCCAAAACCTCC	959
CC	The manufacturing method of DNA polymerase of active hepatitis B virus	300	pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr	320
CC	comprises transforming a host cell simultaneously with an expression	960	TGTAATAGACTATTGATTGGAAAGTATGTCAAAGAAATTTGGGCTCTTTGGGCTTTGC	1019
CC	vector including a coding gene of hepatitis B virus DNA polymerase and an	320	oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl	340
CC	expression vector including a coding gene of GRP94 protein, culturing the	1020	TGCCCCCTTTTACCAATGTGGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1079
CC	transformed host under the condition to initiate replication at two of	340	aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe	360
CC	replication origins of the expression vectors, and purifying DNA	1080	TAAAGAGCTTTTCACTTCTCGCAACTTACAAAGCCCTTTCTGTGTAAACAATATCTGAA	1139
CC	polymerase of hepatitis B virus included in the culture solution. This	360	rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi	380
CC	sequence represents a Hepatitis B virus protein relating to the GRP94	1140	CCTTTTACCCCGTTGCCCGCAACGGTCCGGTCTCTGCAAGTGTGTTGTGACGCAACCCC	1199
CC	protein of the invention.			
XX	Sequence 539 AA;			
SQ				
Alignment Scores:				
Pred. No.:	9, 9e-209	Length:	539	
Score:	2737.50	Matches:	517	
Percent Similarity:	96.86%	Conservative:	7	
Best Local Similarity:	95.56%	Mismatches:	15	
Query Match:	46.34%	Indels:	3	
DB:	7	Gaps:	1	
US-10-761-006A-1 (1-3215) x ADM68212 (1-539)				
QY	1 CTCCCAACATTTCCACAGCTCTGCTAGATCCCGAGGTGAGGGCCCTATATTTCTCGC	60		
DB	1 LeuHisHisIleSerProSerProAlaArgSerGlnSerGlnGlyProIlePheSerSer	20		

```

:|||||
380 sLeuYrProValAlaArg--ArgThrAlaLeuCysGlnValPheAlaAspAlaThrPr 399
1200 CACTGGATGGGCTTGCCCATAGCCATCAGCGCATGGCTGGAACTTTCTGGCTCTCT 1259
399 oThrGlyTriPGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe 419
1260 GCGATCCATCTCGCGAACTCTAGCAGCTTGTTTGCTCGGAGCGGTCTGGAGCAA 1319
419 uProfileHisThrAlaGluLeuLeuAlaLaCysPheAlaArgSerArgSerGlyAlaLy 439
1320 ACTTATCGGAACCGACACTGTGTGTCCTCTCGGAATACACTCTCTTCATGGCT 1379
439 sLeuIleGlyThrAspAsnSerValValLeuSerArgGlySerThrPheProThrPle 459
1380 GCTAGGCTGTGTCGCCAACTGGATCTCTGCGCGGAGCTCTCTTGTCTACGTCGCCGTCGCC 1439
459 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrTyrrPheValTyrrValProSerAl 479
1440 GCTGAATCCCGGAGCAGCCGCTCTCGGGCCGTTTGGGGCTCTACGTCCTCTTCTCA 1499
479 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuIleArgProLeuLeuHi 499
1500 TCTGCCGCTTCCGGCCGACCAACCGGCGCACCTCTCTTTACGCGGTCTCCCGATGTGCC 1559
499 sLeuArgPheArgProThrThrGlyArgThrSerLeuTyrrAlaValSerProSerValPr 519
1560 TTCTCATCTCCGAGACCGTGTGCACTTTCGCTTCACTCTGCACTGCGCATGAGACCA 1619
519 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaIleArgProPr 539
1620 G 1620
539 o 539

RESULT 10
ADX40771
ID ADX40771 standard; protein; 842 AA.
AC
AC ADX40771;
XX
XX 21-APR-2005 (first entry)
XX
XX HBV polymerase protein #14.
XX
XX Immune stimulation; polymerase; enzyme.
XX
XX Hepatitis B virus.
XX
XX WO2005012502-A2.
XX
XX 10-FEB-2005.
XX
XX 29-MAR-2004; 2004WO-US009510.
XX
XX 28-MAR-2003; 2003US-0458026P.
XX
XX (EPIM-) EPIMMUNE INC.
XX
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MU;
XX
XX WPI; 2005-132661/14.
XX
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
XX acids in length comprising primary anchor residues of the same HLA class
XX I binding motif.
XX
XX Disclosure; Page 380-385; 458pp; English.
XX
XX The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of the
XX peptide epitope, comprising identifying, from a particular antigen of an
```

```

CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HBV polymerase protein used in the scope of the invention.
XX
XX Sequence 842 AA;
SQ
```

```

Alignment Scores:
Pred. No.: 1.95e-208 Length: 842
Score: 2734.50 Matches: 517
Percent Similarity: 96.86% Conservative: 7
Best Local Similarity: 95.56% Mismatches: 15
Query Match: 46.29% Indels: 3
DB: Gaps: 1
```

US-10-761-006A-1 (1-3215) x ADX40771 (1-842)

```

QY 1 CTCACAAATTCACCAAGCTCTGCTAGATCCAGGTTGAGGGCGCTATATTTTCTCTGC 60
Db 304 LeuHisIleSerProSerProAlaArgSerGlnSerGluGlyProIlePheSerSer 323
QY 61 TGTGTGCTCCAGTTCGGAACAGTAAACCTGTTCGAGTACTGCTCTCTCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrrCysLeuThrHisIleVal 343
QY 121 AATCTTCTCGAGACTGGGACCCCTGCACCGAATCGAGAGAACACAACTCAGGATTCCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGGACCCCTGCTCGTGTGTACAGCGGGGTTTCTCGTTGACAAGAATCTCTCAATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGAGACCCACGCTGTTCC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
QY 300 CTGGCCAAAATTCGAGTCCCAACCTCCAACTCCTCACTCACAACCTCTTGTCTCTCAATTT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 423
QY 360 GTCTGCTGCTATCGTGATGTCTGCGGGGTTTATCATATTTCTCTCTCATCTCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrrHisIleProLeuHisProAlaAl 443
QY 420 TATGCTCATCTCTTGTGTTGTTCTTCTGAGTCTTACCAAGGTATGTTGCCGTTGTCCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProA-gTyrrValAlaArgLeuSerSe 463
QY 480 TACTTCCAGGAACATCAACACACAGCACGGGGCCATGCAAGACCTCAGACGCTCTGCTTC 539
Db 463 rThrSerArgAsnIleAsnHisGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
QY 540 AAGGAACCTCTAGTTCCTCTGTTGCTGTACAAAACCTTCGGACGGAAACTGACACTT 599
Db 483 rArgAsnLeuTyrrValSerLeuLeuLeuLeuTyrrLysThrPheGlyArgLysLeuHisLe 503
QY 600 GTATTCCTCCATCCCATCATCTCTGGGCTTTCGCAAGATTCTCTATGGAGTGGGCTCAGTCC 659
Db 503 uTyrrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyGlyLeuSerPr 523
QY 660 GTTTCCTCGGCTCAGTTTACTAGTCCCATTTGTTTTCAGTGGTTTCGAGGGCTTCCCCCA 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 543
QY 720 CTGTTTGGCTTTTCAGTTATATGATCATCTGTGGTATTGGGGCGCAAGTCTGTCAACATCT 779
Db 543 sCysLeuAlaPheSerTyrrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
Db 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
```

```
Qy 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACCTTATGCGATATCTAATTTGGAAGTTG 899
Db 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
Qy 900 GGGTACTTTACCGCAGGAACATATTGTAATAAACTCAAGCAATGTTTTCGAAAACGTC 959
Db 603 pGlyThrLeuProGlnGluHisIleValLeuLysIleValLeuLysGlnCysPheArgLysLeuPr 623
Qy 960 TGTAAATAGACTATTGATGGAAGATGTCAAAGAAATTGCGGTCTTTTGGGCTTTGCG 1019
Db 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
Qy 1020 TGCCCTTTTACAAATGTGCTACTCCCTGCTGCTTATATGATGCTATACAAATC 1079
Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
Qy 1080 TAAGCAGGCTTTCACTTTCTCGCAACTTACAAGGCTTTCTGTGTAAACAATATCTGAA 1139
Db 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi 683
Qy 1140 CTTTATCCCTTCCCGCAGCAACGGTCCGGTCTCTGCAAGTGTGTGCTGACGCAACCCC 1199
Db 683 sLeuTyrProValAlaArg---ArgThrAlaLeuCysGlnValPheAlaAspAlaThrPr 702
Qy 1200 CACTGGATGGGCTTGGCCATAGGCCATAGCCGATCGCGCATGCTCGAACCTTTCTGGCTCT 1259
Db 702 oThrGlyTrpGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe 722
Qy 1260 GCCGATCCATCTCGGAACCTCTAGCAGCTTGTGTTCTCGCAGCGGTCTGGAGCAAA 1319
Db 722 uProIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLys 742
Qy 1320 ACTTATCGGAACCGCAACTCTGTGTCTCTCTCGAAATACACCTCTTTTCCATGGCT 1379
Db 742 sLeuIleGlyThrAspAsnSerValAlaLeuSerArgLysTyrThrSerPheProTriple 762
Qy 1380 GCTAGGTTGTGTCGAACCTGATCTCGCGGAGCTCTTGTCTACGTCCTCGCTCGGC 1439
Db 762 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrTyrPheValTyrValProSerAl 782
Qy 1440 GCTGAATCCCGGACGACCGCTCTCGGGCGGCTTGGGCTCTACCGTCCCTCTTTCA 1499
Db 782 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuIleArgProLeuLeuHi 802
Qy 1500 TCTGCGGTTCCGGCCGACACCGCGCGCACCTCTCTTTACGCGGTCTCCCGGTATGTGCC 1559
Db 802 sLeuArgPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 822
Qy 1560 TTCTCATCTGCGGACCGGTGTCATCTCGCTTCACTCTGCACTGCGATGAGACGACC 1619
Db 822 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 842
Qy 1620 G 1620
Db 842 o 842
RESULT 11
ADX40774
ID ADX40774 standard; protein; 825 AA.
XX
AC AC
XX
DT 21-APR-2005 (first entry)
XX
DE HBV polymerase protein #17.
XX
KW Immune stimulation; polymerase; enzyme.
XX
OS Hepatitis B virus.
XX
PN WO2005012502-A2.
XX
```

```
PD 10-FEB-2005.
XX
XX 29-MAR-2004; 2004WO-US009510.
XX
PR 28-MAR-2003; 2003US-0458026P.
XX
XX (SPIM-) EPIMMUNE INC.
XX
PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX
XX WPI; 2005-132661/14.
XX
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
XX acids in length comprising primary anchor residues of the same HLA class
XX I binding motif.
XX
XX Disclosure; Page 380-385; 458pp; English.
XX
XX The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of the
XX peptide epitope, comprising identifying, from a particular antigen of an
XX infectious agent, variants of a peptide epitope comprising primary anchor
XX residues of the same HLA class I binding motif. The method is useful for
XX identifying a candidate peptide epitope, which induces an HLA class I CTL
XX response against variants of the peptide epitope. This sequence
XX represents an HBV polymerase protein used in the scope of the invention.
XX
XX Sequence 825 AA;
SQ
Alignment Scores:
Pred No.: 9,76e-205 Length: 825
Score: 2688.00 Matches: 504
Percent Similarity: 95.92% Conservative: 13
Best Local Similarity: 93.51% Mismatches: 21
Query Match: 45.51% Indels: 2
DB: Gaps: 0
US-10-761-006a-1 (1-3215) x ADX40774 (1-825)
Qy 7 AACATTCCACCAAGCTCTGTAGATCCAGGGTGAGGGCCCTATATTTCTGCTGGTGG 66
Db 288 HisLeuSerThrThrLysArgGlnSerSerSerGlyHisAlaValGluThrCysTrpTr 307
Qy 67 CTCAGATTCCGGAACAGTAAACCTGTTCCGACTACTGCTCTCCCATATGCTCAATCTT 126
Db 308 LeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAsnLeu 327
Qy 127 CTCGAGGACTGGGACCTCGACCGAACATGAGAGAACACATCAGGATTCCTAGGACC 186
Db 328 LeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArgThr 347
Qy 187 CTTGCTCTGTGTACAGCGGGGTTTTCTCGTTGACAAGATCTTCACAAATACCCAGAG 246
Db 348 ProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThrThrGlu 367
Qy 247 TCTAGACTCTG- GTGGACTCTCTCAATTTCTAGGGGAGCACCACGCTGTTCTGTGCC 305
Db 368 Ser***LeuValValAspPheSerGlnPheSerArgGlySerThrGlnVal-SerTrpPr 387
Qy 306 AAAATTTCGAGTCCCAACCTCCAAATCACTACCAACCTCTTGTCTCTCAATTTGTCCTG 365
Db 387 oLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSerTr 407
Qy 366 GCTATCGCTGGATGTCTCGGGCGTTTTTATCATATTTCTCTCTCTCTCTCTCTCTATGCC 425
Db 407 pLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAlaMetPr 427
Qy 426 TCATCTTCTTGTGTTCTTCTGGACTACCAAGGTATGTTGCCGCTTCTCTCTCTCTCTCTC 485
Db 427 oHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerThrSe 447
Qy 486 CAGGAACATCAACACCCAGCACGGGGCCATGCAAGACCTGACGACTCTCTGCTCAAGGAA 545
```

```
Db 447 rArgAsnIleAsn***GlnHisGlyThrMetGlnAspLeuHisAspSerCysSerIysHis 467
546 ACTTACTGTTCCCTCTCTGTGTGTACAAAACCTTCGGACGGAACCTGCACCTGTATTTC 605
467 sLeuTyrValSerLeuLeuLeuLeuTyrIysThrPheGlyArgIysLeuHisLeuTyrSe 487
606 CCATCCCATCATCTCTGGGCTTTTCGGAAGATTCCTATGGGAGTGGGCTCAGTCCGTTTCT 665
487 rHisProIle***LeuGlyPheArgIysIleProMetGlyValGlyLeuSerProPheLe 507
666 CTTGGCTCAGTTACTAGTCCATTTGTTTCAGTGGTTTCGTAGGCTTTCCCCACACTGTTT 725
507 uLeuAlaGlnPheThrSerIleAlaIleCysSerValValArgAlaPheProHisCysLe 527
726 GGTCTTTCAGTTATATGATGATGTGTATTTGGGGGGAAGTCTGTACAAATCTTGAGTC 785
527 uAlaPheSerTyrMetAspAspValValLeuGlyAlaIysSerValGlnHisLeuGluAl 547
786 CTTTTTACCTCTATTACCAATTTCTTTTGTCTTTTGGGTATACATTTAAACCTTAATAA 845
547 aLeuTyrThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAsnLy 567
846 AACCAACGTTGGGCTACTCCCTTAACCTTCATGGGATATGTAATTGGAAGTTGGGGTAC 905
567 sThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrpGlyTh 587
906 TTTTACGACGAGACATATTGTACTAAACCTCAAGCAATGTTTTCGAAAACGTGCTGTAAA 965
587 rLeuProGlnAspHisIleValGlnIysLeuIysGlnCysPheArgIysLeuProValAs 607
966 TAGACCTATTGATTGAAAGTATGTCAAGAATGTGGGTCTTTTGGCTTCTGCTGCCCC 1025
607 nArgProIleAspTrpIysValCysGlnArgIleValGlyLeuLeuGlyPheAlaIaPr 627
1026 TTTTACACATGTGGCTATCTCGCTCTGATGCTCTTTATATGATGATATCAATCTAAGCA 1085
627 oPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerIysGl 647
1086 GGCTTTACCTTTCTCGCAACTTCAAGGCTTTCTGTGTAAACAATATCTGAACCTTTA 1145
647 nAlaPheThrPheSerProThrTyrIysAlaPheLeuCysIysGlnTyrLeuAsnLeuTy 667
1146 CCCCCTTCCCGGCAAGCTCTCGGCTCTCTCGAAGTGTCTGACGCAACCCCACTGG 1205
667 rProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProThrGl 687
1206 ATGGGCTTGGCCATAGGCCATCAGCGCATGGCTGGAACCTTTCTGCTCTCTGCGCAT 1265
687 yTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLeuProIl 707
1266 CCATCTCGCGAACTCTCTAGCAGCTTGTGTTGCTCGACGCGGTCTGGAGCAAACTTAT 1325
707 eHisThrAlaGluLeuLeuAlaIaCysPheAlaArgSerArgSerGlyAlaIysLeuIl 727
1326 CGGAACCGCAACTCTGTGTCTCTCTCGAAATATACCTCTCTTCCATGGCTGTAGG 1385
727 eGlyThrAspAsnSerValValLeuSerArgIysTyrThrSerPheProTrpLeuLeuGl 747
1386 GTCTGTCTGCCAATCTGGATCTCGCGGAGCTCTTTGTCTAGTCCCGTCCGCGCGCTGAA 1445
747 yCysAlaIaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAlaLeuAs 767
1446 TCCCGGCGAGACCCGCTCTCGGGCGGCTTTGGGGCTCTACCGTCCCTCTCTTCATCTGCC 1505
767 nProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHisLeuPr 787
1506 GTTCCGGCGACACGCGGGGCGACTCTCTTTTAGCGGTCTCCCGTAGTGTCTTCTCTCA 1565
787 oPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValProSerHi 807
1566 TCTGCGGCGCGGTGACACTTCTCCTCTGACGTCGATGGAGACCCG 1620
```

```
Db 807 sLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 825
RESULT 12
ADX40773
ID ADX40773 standard; protein; 827 AA.
XX
AC ADX40773;
XX
DT 21-APR-2005 (first entry)
XX
DE HBV polymerase protein #16.
XX
KW Immune stimulation; polymerase; enzyme.
XX
OS Hepatitis B virus.
XX
PN WO2005012502-A2.
XX
PD 10-FEB-2005.
XX
PF 29-MAR-2004; 2004WO-US009510.
XX
PR 28-MAR-2003; 2003US-0458026P.
XX
PA (SPIM-) EPIMUNE INC.
XX
PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
WPI; 2005-132661/14.
XX
PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX
PS Disclosure; Page 380-385; 458pp; English.
XX
CC The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HBV polymerase protein used in the scope of the invention.
SQ
Sequence 827 AA;
Alignment Scores:
Pred. No.: 5.57e-204 Length: 827
Score: 2678.50 Matches: 504
Percent Similarity: 95.91% Conservative: 12
Best Local Similarity: 93.68% Mismatches: 18
Query Match: 45.34% Indels: 5
DB: Gaps: 1
US-10-761-006A-1 (1-3215) x ADX40773 (1-827)
QY 19 AGCTCTGCTAGATCCAGGGGTGAGGG-----CCTATATTTTCTGCTGGTGGCTC 69
Db 291 ThrSerLysArgGlnSerSerSerGlyHisAlaValGluLeuHisProCysTrpTrpLeu 310
QY 70 CAGTTCCGGAACAGTAAACCTGTTCGACTACTGCTCTCCCATATCGTCAATCTTCTC 129
Db 311 GlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAsnLeuLeu 330
QY 130 GAGGACTGGGGCCCTGCGACCAACATGAGACACCAACATCAGGATTCCTAGGACCCCT 189
Db 331 GluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArgThrPro 350
QY 190 GCTCGTGTTCACAGGGGGGGTTCCTGTTTGACAAAGATCCTCACAATACCGCAGAGTCT 249
Db 351 AlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrThrGluSer 370
```

QY 250 AGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGAGCACCACGCTGTTCTCGGCGAAA 308
DB |||||
QY 371 ArgLeuValValAspPheSerGlnPheSerArgGlySerThr**Val-SerTrpProLy 390
DB |||||
QY 309 ATTCGCAGTCCCAACTCCAACTCACACCACTCTCTGCTCTCCAAATTTGCTCGCT 368
DB |||||
QY 390 sPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSerTrpLe 410
DB |||||
QY 369 ATCCCTCGATGTGCTCGGGGTTTTATCATATTCCTCTTCATCCTCTGCTATGCTCTCA 428
DB |||||
QY 410 uSerLeuAspValSerAlaAlaPheTyHieIleProLeuHisProAlaAlaMetProHi 430
DB |||||
QY 429 TCTTCTTGTGTTCTCTGACCTACCAAGTATGTTGCCGTTGCTCTCTACTCTCCAG 488
DB |||||
QY 430 sLeuLeuValGlySerSerGlyLeuProArgTyValAlaArgLeuSerSerThrSerAr 450
DB |||||
QY 489 GAACATCAACACACGACGAGGGGCATGCAAGACTCTGACGACTCTCTGCTCAAGGAACT 548
DB |||||
QY 450 gAsn**AsnTyHieIleGlnHisGlyThrMetGlnAspLeuHisGluSerCysSerArgAsnLe 470
DB |||||
QY 549 CTACGTTTCCCTCTGTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 608
DB |||||
QY 470 uTyHieIleSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 490
DB |||||
QY 609 TCCCATCATCTCGGCTTCCGACGATTCCTATGGAGTGGGCTCAGTCCGTTCTCTCT 668
DB |||||
QY 490 sProIleIleLeuGlyPheArgGlyIleProMetGlyValGlyLeuSerProPheLeuLe 510
DB |||||
QY 669 GGCTCAGTTTACTAGTGCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 728
DB |||||
QY 510 uAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHisCysLeuAl 530
DB |||||
QY 729 TTTTACGTTATATGATGATGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 788
DB |||||
QY 530 aPheSerTyMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGluSerLe 550
DB |||||
QY 789 TTTTACCTCTATACCAATTTCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 848
DB |||||
QY 550 uPheThrAlaIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProHisLysTh 570
DB |||||
QY 849 CAACGTTGGGGCTACTCCCTTAATCTCATGGATATGTAATTAATGGAAGTTGGGCTACTTT 908
DB |||||
QY 570 rLyArgTrpGlyTySerLeuAsnPheMetGlyTyHieIleGlySerTrpGlyThrLe 590
DB |||||
QY 909 ACCGAGGAACATATTTGTTACTAAACTCAAGCAATGTTTTCGAAACTGCTGTAAATAG 968
DB |||||
QY 590 uProGlnGluHisIleValGlnLysIleLysGlnCysPheArgLysLeuProValAsnAr 610
DB |||||
QY 969 ACCTATTGATTGGAAAGTATGTCAGAAAGTATGTTGGGCTTTTGGGCTTTGCTGCGCTTT 1028
DB |||||
QY 610 gProIleAspTrpLysValCysGln**IleValGlyLeuLeuGlyPheAlaAlaProPh 630
DB |||||
QY 1029 TACAAATGTCGCTATCTGCTGCTGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 1088
DB |||||
QY 630 eThrGlnCysGlyTyProAlaLeuMetProLeuTyHieIleGlnAlaLysGlnAl 650
DB |||||
QY 1089 TTTTCACTTTCTCGCAACTTCAAGGCTTTCTGTTGTTAAACAATATCTGAACCTTTACCC 1148
DB |||||
QY 650 aPheThrPheSerProThrTyHieIlePheLeuCysLysGlnTyHieIleLeuHisLeuTyPr 670
DB |||||
QY 1149 GCTTCCCGGCGAACGGTCCGCTCTCTGCAAGTGTTCCTGACGCAACCCCTGCTGATG 1208
DB |||||
QY 670 oValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProThrGlyTr 690
DB |||||
QY 1209 GGGCTTGGCCATAGCCATCAGCGATGCTGGAACCTTTCTGGCTCTCTGCTGCGATCCA 1268
DB |||||
QY 690 pGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLeuProIleHi 710
DB |||||
QY 1269 TACTCGGAACCTCTAGCAGCTTGTGTTGCTGCGACGGCTCTGGAGCAAACTTATCGG 1328
DB |||||
QY 710 eThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLysLeuIleGl 730
DB |||||

QY 1329 AACCGAACACTCTGTTGTTCTCTCTCGAAATACACCTCCTTCCATGGCTGTAGGGTG 1388
DB |||||
QY 730 yThrAspAsnSerValValLeuSerArgLySlyThrSerPheProTrpLeuLeuGlyCy 750
DB |||||
QY 1389 TGTGTCAACTGATCCTCGCGGAGCGTCTTGTCTAGTCCGCTCGGCGCTGAATCC 1448
DB |||||
QY 750 sAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyHieIleProSerAlaLeuAsnPr 770
DB |||||
QY 1449 CGCGGAGCAGCCCTCTCGGGGCGGTTTGGGGCTCTACCGTCCCTCTTCTCATCTGCCGCT 1508
DB |||||
QY 770 oAlaAspAspProSerArgGlyArgLeuGlyLeuTyHieIleProLeuLeuHisLeuProPh 790
DB |||||
QY 1509 CCGGCCGACACCGGCGGCGACCTCTCTTTACGGCGTCTCCCGTATGTGCTTCTCATCT 1568
DB |||||
QY 790 eArgProThrThrGlyArgThrSerLeuTyHieIleValSerProSerValProSerHisLe 810
DB |||||
QY 1569 GCGGAGCCGTTGTCACCTCTGACGCTGCGATGGAGACCCG 1620
DB |||||
QY 810 uProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 827
DB |||||
RESULT 13
AAE04708
ID AAE04708 standard; protein; 842 AA.
XX AC
XX AAE04708;
XX
XX 04-SEP-2001 (first entry)
XX
XX Hepatitis B virus FRI strain genotype G HBpol protein.
XX
XX HBV genotype G; precore; HBpol; polymerase; envelope protein; preS1;
XX KW preS2; surface antigen; HBsAg; HBx protein; vaccine; liver disease;
XX KW hepatitis; liver cancer; HBcAg; core antigen.
XX
XX Hepatitis B virus.
XX OS
XX WO200138498-A2.
XX
XX 31-MAY-2001.
XX
XX 21-NOV-2000; 2000WO-US032108.
XX PF
XX 24-NOV-1999; 99US-0167206P.
XX PR
XX (PHAR-) PHARMASSET INC.
XX PA (INNO-) INNOGENETICS NV.
XX
XX Stuyver L, Schinazi R, De Gendt S, Van Geyt C, Zoulim F, Fried M,
XX PI Rossau R;
XX
XX WPI; 2001-367676/38.
XX DR N-PSDB; AAD09091.
XX
XX Novel hepatitis B virus genotype G, nucleic acids encoding virus,
XX PT polypeptides encoded by nucleic acids, useful for preparing vaccine to
XX PT treat or prevent the hepatitis B virus genotype G infection in a subject.
XX
XX Example; Fig 6; 84pp; English.
XX
XX The present invention relates to hepatitis B virus (HBV) strain FRI,
XX CC genotype G DNA encoding PreCore/Core protein, HBpol, envelope (preS1,
XX CC preS2 and surface antigen HBsAg) and HBx proteins. HBV genotype G nucleic
XX CC acids and polypeptides are useful for diagnosing, prognosing and treating
XX CC infections caused by HBV genotype G. They can be used in a vaccine to
XX CC treat or prevent HBV genotype G infection. The HBV genotype G derived
XX CC nucleic acids and antibodies are useful for detecting HBV genotype G in a
XX CC sample or diagnosis of HBV genotype G infection. The presence of HBV
XX CC genotype G statistically correlates with the presence of liver damage
XX CC and/or liver cancer in the subject. The HBV genotype G core insert
XX CC peptide encoding nucleic acid is useful for designing monitoring assays
XX CC to study and predict the evolution of anti-HBe and anti-HBc antibodies
XX CC and HBsAg (genotype G e antigen) in patients infected with HBV. The
XX CC antibodies or antigens of HBV genotype G are useful for identifying a

CC stage of liver disease caused by HBV genotype G. The present sequence is
CC hepatitis B virus (HBV) strain FRI, genotype G polymerase protein, HBpol
XX
SQ Sequence 842 AA;

Alignment Scores:
Pred. No.: 1,066-203 Length: 842
Score: 2675.00 Matches: 496
Percent Similarity: 96.30% Conservative: 25
Best Local Similarity: 91.68% Mismatches: 19
Query Match: 45.29% Indels: 2
DB: 4 Gaps: 0

US-10-761-006A-1 (1-3215) x AAE04708 (1-842)

QY	1	CTCCACAACTCCACCAAGCTCTGCTAGATCCAGGGTGAGGGGCTATATTTCTCTGC	60
DB	303	LeuTyr:SerileProPheSerThrLysSerGlnSerGlnGlyProValPheSerCys	322
QY	61	TGTTGGCTCCAGTTCGGGAACAGTAACCTGTTCCGACTACTGCTCTCCCATATCGTC	120
DB	323	TrpTrpLeuGlnPheArgAspSerGluProCysSerAspTyrCysLeuSerHisLeuVal	342
QY	121	AATCTTCTCGAGACTGGGGACCTCGACCGAACATGGAGAACACACATCAGGATTCCT	180
DB	343	AsnLeuLeuGlnAspTrpGlyProCysThrGluHisGlyGluHisIleArgIlePro	362
QY	181	AGGACCCCTGCTCGTGTACAGCGGGGTTTTCTCGTTGACAAAGAACTCTCAATACC	240
DB	363	ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr	382
QY	241	GCAGAGTCTAGACTCTG-GTGGACTTCTCAATTTCTAGGGGGAGACCCACAGTGTTC	299
DB	383	AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerAlaArgVal-Se	402
QY	300	CTGGCCAAATTCGAGTCCCAACTCCAACTCACTCAACCACTCTTGTCTCCCAATTT	359
DB	402	rTTPpPolysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe	422
QY	360	GTCTGGCTATCGCTGGATGTCTGGCGGTTTTATCATATTCCTCTTCATCTCTCTGC	419
DB	422	uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl	442
QY	420	TATGCTCATCTTCTGTGGTCTTCTGCACTACCAAGTATGTGGCGTTTGCTTC	479
DB	442	aMetProHisLeuLeuValGlySerGlyLeuSerArgTyrValAlaArgLeuSerSe	462
QY	480	TACTTCCAGGAACATCAACACGACGACGGGCCATGCAAGACTCGCAGACTCTCTGCTC	539
DB	462	rAspSerArgIleLeuAspHisGlnTyrGlyThrLeuGlnAsnLeuHisAspSerCysSe	482
QY	540	AAGAAACTCTAGCTTTCCCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	599
DB	482	rArgGlnLeuTyrValSerLeuMetLeuLeuTyrLysThrPheGlyArgLysLeuHisLe	502
QY	600	GTATTCATCCCATCATCTCGGCTTTCGCAAGATTCCTATGGAGTGAGGCTCTAGTCC	659
DB	502	uTyrSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr	522
QY	660	GTCTCTCTCGCTCAGTTACTAGTGCATTTCTGTCAGTGTCTGTTAGGCTTTCCCCCA	719
DB	522	oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi	542
QY	720	CTGTTTGGCTTTTCAGTTATATGATGATGTGTATTGGGGCGAAGTCTGTACAACTCT	779
DB	542	sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe	562
QY	780	TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAACCC	839
DB	562	uGluSerLeuTyrThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr	582
QY	840	TAATAAAACCAACGTTGGGGCTACTCCCTTAACCTCATGGGATATGTAATTCGAAAGTTG	899

DB	582	oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr	602
QY	900	GGGTACTTTACCGCAGGAACATATTGTACTAAACTCAAGCAATGTTTTCGAAAACTGCC	959
DB	602	pGlyThrLeuProGlnGluHisIleThrGlnLysIleLysGlnCysPheArgLysLeuPr	622
QY	960	TGTAATATAGACCTATTGATTGGAAAGTATGTCAAGAATTGGGTCTTTTGGCTTTGC	1019
DB	622	oValAsnArgProIleAspTrpLysValCysGlnArgIleThrGlyLeuLeuGlyPheAl	642
QY	1020	TGCCCCCTTTTACACAATGTGGCTATCTCGCTTGATGCTCTTTATATGATGATATCAATC	1079
DB	642	alaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnAl	662
QY	1080	TAAGCAGGCTTTCACTTTCTGCCAACTTCAAGGCTTTCTGTGTAAACAATATCTGAA	1139
DB	662	alysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrMetAs	682
QY	1140	CCTTTACCCCGTTGCCCGCAACGGTCCGCTCTCTGCCAAGTGTCTGCTGACGCAACCCC	1199
DB	682	nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr	702
QY	1200	CACTGGATGGGCTTGGCCATAGGCCATCAGCGCATGGCTGGAACTTTCTGCTCTCTCT	1259
DB	702	oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe	722
QY	1260	GCGATTCATATCGCGAACTCTCTAGCAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGT	1319
DB	722	uProIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys	742
QY	1320	ACTTATCGGAACCGCAACTCTGTGTCTCTCTCGGAATACACCTCTCTTCCATGGCT	1379
DB	742	sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProIlePr	762
QY	1380	GCTAGGCTGTGTGCCAACTGGATCTCGCGCGGAGCTCTTGTCTACGTCCTCGCGC	1439
DB	762	uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl	782
QY	1440	GCTGAATCCCGGAGACGACCCGCTCTCGGGCGGTTTGGGGCTCTACCGTCCCTTCTTCA	1499
DB	782	aLeuAsnProAlaAspProSerArgGlyArgLeuGlyLeuCysArgProLeuLeuAr	802
QY	1500	TCTGCGGTTCCGGCGGACCGACCGGCGCACCTCTCTTTACGGGCTCTCCCGTATGTGC	1559
DB	802	gLeuProPheLeuProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr	822
QY	1560	TTCTCATCTCGCGGACCGGTGTGCATCTTCCTCTTCACCTCTGCACGTGCGATGGAGACC	1619
DB	822	oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValThrTrpLysProPr	842
QY	1620	G 1620	
DB	842	O 842	
RESULT 14			
AAAG6918	AAG66918 standard; protein; 842 AA.		
XX	AAG66918;		
AC	19-OCT-2001 (first entry)		
DT	HBV pres1/52/HBS region protein.		
XX	Hepatitis B virus; HBV; preCore; Core; pres1; pres2; HBS; HBX; HBPol;		
KW	HBSag; antiviral; vaccine; genotype G; genotyping; HBSag; HBeAg.		
XX	Hepatitis B virus.		
XX	W0200140279-A2.		
XX	07-JUN-2001.		

PF 20-NOV-2000; 2000WO-EP011526.
 XX
 PR 03-DEC-1999; 99EP-00870252.
 PR 07-DEC-1999; 99US-0169287P.
 XX
 PA (INNO-) INNOGENETICS NV.
 XX
 PI Stuyver L, Van Geyt C, De Gendt S;
 XX
 DR WPI; 2001-374785/39.
 DR N-PSDB; AAH77562.
 XX
 PT Novel isolated and/or purified hepatitis B virus polypeptide and
 PT polynucleotide sequences that are phylogenetically different from HBV
 PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
 PT therapy.
 XX
 PS Claim 16; Fig 1; 94pp; English.
 XX
 CC The invention relates to the complete nucleic acid sequence of a new
 CC human hepatitis B virus (HBV) genotype, provisionally named genotype G.
 CC This genotype was found with a high prevalence in patients chronically
 CC infected with HBV and residing in Europe and the USA. The invention
 CC relates to a fully defined sequence of 3248 nucleotides as given in
 CC specification, a sequence with 92% identity to the given sequence, or
 CC sequence that is degenerate to the mentioned sequences. These
 CC polynucleotides are useful for HBV genotyping. The proteins encoded by
 CC the polynucleotides are useful for detecting antibodies in a biological
 CC sample. Ligands that bind to the proteins and antibodies directed against
 CC the proteins are useful for detecting the proteins and for detecting
 CC HBsAg and HBeAg (precursor proteins). They are also useful for
 CC preparing a vaccine or medicament for treating HBV infections. The
 CC present sequence is encoded by the genome of HBV genotype G strain FRI
 XX
 SQ Sequence 842 AA;
 Alignment Scores:
 Pred. No.: 1,06e-203 Length: 842
 Score: 2675.00 Matches: 496
 Percent Similarity: 96.30% Conservative: 25
 Best Local Similarity: 91.68% Mismatches: 19
 Query Match: 45.23% Indels: 2
 DB: 4 Gaps: 0
 US-10-761-006a-1 (1-3215) x AAG66918 (1-842)
 QY 1 CTCACAACTTCCACCAAGCTCTCTAGATCCCAAGGCTGAGGGCCATATTTTCCTGC 60
 DB 303 LeuTyrSerIleProProSerSerThrIysSerGlnSerGlnGlyProValPheSerCys 322
 QY 61 TGGTGGCTCCAGTTCGGAACAGTAACCTGTTCGACTACTGCTCTCCGATATCGTC 120
 DB 323 TrpTrpLeuGlnPheArgAspSerGluProCysSerAspTyrCysLeuSerHisLeuVal 342
 QY 121 AATCTTCCGAGGACTGGGACCTGCACCGAACATGAGAACACAAATCAGGATTCCT 180
 DB 343 AsnLeuLeuGlnAspTrpGlyProCysThrGluHisGlyGluHisIleArgIlePro 362
 QY 181 AGGACCTCTGCTGCTGACGGGGGTTTCTCTGTTGACAAAGATCCTCACAAATACC 240
 DB 363 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 382
 QY 241 GCAGAGCTTAGACTCTCTCAATTTCTAGGGGGAGCACCCACGCTGTC 299
 DB 383 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerAlaArgVal-Se 402
 QY 300 CTGGCCAAATTCGAGTCCCACTCCCAATCACTCACCAACCTTGTCTCTCCCAATT 359
 DB 402 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 422
 QY 360 GTCTCGCTATCGCTGATGTCTGCGGCGGTTTATCATATTCCTCTTCATCTCTGCTGC 419
 DB 422 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 442

QY 420 TATGCCTCATCTTCTTGTGGTTCTTCTGACTACCAAGGTATGTGCCGTTGTCTCTC 479
 DB 442 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 462
 QY 480 TACTTCAGAAACATCAACACGACGCGGGCCATGCAAGACCTGCACGACCTCTCTGCTC 539
 DB 462 rAspSerArgIleLeuAspHisGlnTyrGlyThrLeuGlnAsnLeuHisAspSerCysSe 482
 QY 540 AAGGAACTCTAGCTTTCCTCTTGTGTGTACAAAACCTTCGGACGGAACCTGCACATT 599
 DB 482 rArgGlnLeuTyrValSerLeuMetLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 502
 QY 600 GTATTCCTCCATCCATCATCTCGGCTTTCGCAAGATTCTTATGGAGTGGGCTCTAGTCC 659
 DB 502 uTyrSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 522
 QY 660 GTTTCCTCGCTCAGTTACTAGTGCATTTGTTCAGTGGTTCGTAGGCTTCCCTCCA 719
 DB 522 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHI 542
 QY 720 CTGTTGGCTTTCAGTTATATGATGATGTGTTATGGGGCGCAAGTCTGTACAACATCT 779
 DB 542 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 562
 QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTTCTTGTCTTGGGTATACATTTAAACCC 839
 DB 562 uGluSerLeuTyrThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 582
 QY 840 TAATAAACCAACGTTGGGCTACTCCCTTAACCTTACCTGATGATGATGATGATGATG 899
 DB 582 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 602
 QY 900 GGTGTACTTACCAGGAGACATATTGACTAAACCTCAAGCAATGTTTTCGAAACCTGCC 959
 DB 602 pGlyThrLeuProGlnGluHisIleThrGlnLysIleLysGlnCysPheArgLysLeuPr 622
 QY 960 TGTAAATAGACCTATTGATTGGAAGTATGTCAAAGAAATTGTGGTCTTTTGGGCTTTCG 1019
 DB 622 oValAsnArgProIleAspTrpLysValCysGlnArgIleThrGlyLeuLeuGlyPheAl 642
 QY 1020 TGCCCTTTTACAAATGTCCTATCTGCTGATGCTGCTTATGATGATGATGATGATGATC 1079
 DB 642 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnAl 662
 QY 1080 TAAGCAGCTTCACTTTCTCGCAACTTCAAGGCTTTCTGTGTAACAATATCTGAA 1139
 DB 662 aLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrMetAs 682
 QY 1140 CTTTACCCGCTTCCCGCAACCGTCCGCTCTCTGCAAGTGTTCGTCGACGCAACCCC 1199
 DB 682 nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 702
 QY 1200 CACTGGATGGGCTTGGCCATAGCCATACGCGATGCGGATGCGTGGACCTTCTGGCTCTC 1259
 DB 702 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 722
 QY 1260 GCCGATCCATCTCGGAACCTCTAGCAGCTTGTGTCGACGCGCTTGTGTCGACGAGCAA 1319
 DB 722 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 742
 QY 1320 ACTTATCGAAACCGCAACTCTGTTGTCTCTCTCGGAAATACACCTCTTTCATGCTGCT 1379
 DB 742 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 762
 QY 1380 GCTAGGCTGCTGCCAATCTGATCTCGCGGGGACGCTCTTGTGCTACGCTCCGCTCGGC 1439
 DB 762 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 782
 QY 1440 GCTGATCCCGGACGACCGCTCTCGGGGCGGTTTGGGCTCTACCGCTCCCTCTTCTCA 1499
 DB 782 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuCysArgProLeuLeuLeu 802

Qy	1500	TC	TGCGCGTTCGGCGCGACACGCGGGCGACCTCTCTTTTACGCGGTCTCCCGGTATGTC	1559
Db	802	gLeu	ProPheLeuProThrThrGlyAArgThrSerLeuTyAlaValSerProSerValPr	822
Qy	1560	TT	CTCACTCTCCGCGACCGTGTGCACTTCGCTTCACCTCTGCACGTCCATGGAGACCC	1619
Db	822	oSer	HisLeuProAspArgValHisPheAlaSerProLeuHisValThrTrpLysProPr	842
Qy	1620	G	1620	
Db	842	o	842	
RESULT 15				
ID	AD	X40761	standard; protein; 832 AA.	
XX	AC	AD	X40761;	
XX	DT	21-APR-2005	(first entry)	
XX	DE	HBV	polymerase protein #4.	
XX	KW	Immune	stimulation; polymerase; enzyme.	
XX	OS	Hepatitis B	virus.	
XX	PN	W02005012502-A2.		
XX	PD	10-FEB-2005.		
XX	PF	29-MAR-2004; 2004WO-US009510.		
XX	PR	28-MAR-2003; 2003US-0458026P.		
XX	PA	(EPIM-) EPIMUNE INC.		
XX	PI	Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;		
XX	DR	WPI; 2005-132661/14.		
XX	PT	Identifying a candidate peptide epitope, which induces a HLA class I CTL		
PT	response	comprises identifying variants of a peptide epitope 8-11 amino		
PT	acids	in length comprising primary anchor residues of the same HLA class		
PT	I	binding motif.		
XX	PS	Disclosure; Page 380-385; 458pp; English.		
XX	CC	The invention relates to a method of identifying a candidate peptide		
CC	epitope	which induces an HLA class I CTL response against variants of the		
CC	peptide	epitope, comprising identifying, from a particular antigen of an		
CC	infectious	agent, variants of a peptide epitope comprising primary anchor		
CC	residues	of the same HLA class I binding motif. The method is useful for		
CC	identifying	a candidate peptide epitope, which induces an HLA class I CTL		
CC	response	against variants of the peptide epitope. This sequence		
CC	represents	an HBV polymerase protein used in the scope of the invention.		
XX	SQ	Sequence 832 AA;		
Alignment Scores:				
Pred. No.:	4.97e-202	Length:	832	
Score:	2654.00	Matches:	492	
Percent Similarity:	95.56%	Conservative:	25	
Best Local Similarity:	90.94%	Mismatches:	23	
Query Match:	44.93%	Indels:	2	
DB:	9	Gaps:	0	
US-10-761-006A-1 (1-3215) x ADX40761 (1-832)				
Qy	1	CTCCACACATTCACCAAGCTCTGCTAGATCCAGGCTGAGGGCCCTATATTTTCCTGC	60	
Db	293	LeuHisAenPheProProAsnSerAlaArgSerGlnGlyGluArgProValPheProCys	312	
Qy	61	TGTTGGCTTCAGTTCCGGAACAGTAACCCCTGTGTTCCGACTACTGCCTCTCCCATATGCTC	120	

313	Db	 TtpTrpLeuGlnPheArgAnSerLysProCysSerAspTyrCysLeuSerHisIleVal	332
121	QY	 AATCTTCTCGAGACATGGGGACCTCCACCGAACTATGGAGAACACAAACATCAGGATTCTCT	180
333	Db	 AnuLeuLeuGluAnpTrpGlyProCysThrGluHisGlyGluHisIleArgIlePro	352
181	QY	 AGGACCCCTCGCTCGTGTATACAGCGGGGTTTTCTCGTTGACAAAGAAATCCTCACAATACC	240
353	Db	 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr	372
241	QY	 GCAGACTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGGAGACCCACAGTGTTC	299
373	Db	 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnHisArgVal-Se	392
300	QY	 CTGGCCAAATAATCGCAGTCCCACACCTCCAATCAGTCCACCAACCTCTGTGCTCTCCAAATT	359
392	Db	 rTppProLysPheAlaValProAnuLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe	412
360	QY	 GTCTCGGCTATCGCTGGATGTGTCTCGGGGGTTTTATATATTCCTTCATCCTCTGCTGC	419
412	Db	 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaSe	432
420	QY	 TATGCTCATCTCTTGTGTGTTCTTCTCGACTACCAAGTATGTTGCCGTTGTGCTTC	479
432	Db	 rMetProHisLeuLeuValGlySerThrGlyLeuSerArgTyrValAlaArgValSerSe	452
480	QY	 TACTTCCAGGAACATCAACACCCAGCAGCGGGCCATGCCAAGACCTCCAGCACCTCTGCTC	539
452	Db	 rAnSerArgIlePheAsnHisGlnArgGlyThrMetGlnAsnLeuHisAspTyrCysSe	472
540	QY	 AAGGAAACTCTACGTTTCCCTCTTGTGTGTTGTCATCAAAACCTTCGAGCGGAACTGCAC	599
472	Db	 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrGlnThrPheGlyArgLysLeuHisLe	492
600	QY	 GTATCCCATCCCATCATCTCTGGGCTTTCGCAAGATTCCTATGGGAGTGGGCTCAGTCC	659
492	Db	 uTyrSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr	512
660	QY	 GTTTCTCTCGGCTCAGTTTACTAGTCGCCATTTGTTTCAGTGGTTCGTAGGCTTTC	719
512	Db	 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHI	532
720	QY	 CTGTTTGGCTTTCAGTTATATGGATGATGTGGTATTGGGGGGCAAGTCTGTACAAATCT	779
532	Db	 bCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe	552
780	QY	 TGAGTCCCTTTTACCTCTATTACGAATTTCTTTTGTCTTTGGGGTATACATTTAAACCC	839
552	Db	 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr	572
840	QY	 TAATAAAACCAACGTTGGGGCTACCTCCCTTAACTTTCATGGGATATGTAATTCGAAGTTG	899
572	Db	 oAnuLysThrLysArgTrpGlyTyrSerLeuHisPheMetGlyTyrValIleGlySerTy	592
900	QY	 GGGTACTTTTACCGCAGGAACATATTGTACTATAAACCTCAAGCAATGTTTTCGAAAACTGCC	959
592	Db	 rGlySerLeuProGlnAspHisIleIleGlnLysIleLysGluCysPheArgLysLeuPr	612
960	QY	 TGTAATATAGACCTATTGATTGGAAAGATGTCAAGAATTGCGGGTCTTTTGGGCTTTGCG	1019
612	Db	 oIleAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl	632
1020	QY	 TGCGCCCTTTTACACATGTGGCTATCCTCGCTTGATGCTCTTATATGCAATGTATACAATC	1079
632	Db	 aAlaProPheThrGlnCysGlyTyrProAlaLeuLeuMetProLeuTyrAlaCysIleGlnSe	652
1080	QY	 TAAGCAGGCTTTTCACTTTCTCGCAACTTCAAGGCTTTTCTGTGTAACAAATATCTGAA	1139
652	Db	 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs	672
1140	QY	 CTTTTACCGCGTTGCCGGCAACGGTCCGGTCTCTGCCAAGTGTTCGTGACCGCAACCCC	1199


```

Db      672 nLeuTyrProValalaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 692
Qy      1200 CACTGGATGGGCTTGGCCATAGGCCATAGCGCATGGCTGGAAACCTTTCTGGCTCTCT 1259
Db      692 oThrGlyTrpGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheLeuAlaProLe 712
Qy      1260 GCCGATCCATACCTCGGAACCTCCTAGCAGCTTGTGTCGCGAGCGGCTCTGGAGCAAA 1319
Db      712 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAs 732
Qy      1320 ACTTATCGAACCGACAACTCTGTGTCCTCTCTCGGAAATACACCTCCTTTCCATGGCT 1379
Db      732 nIleLeuGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 752
Qy      1380 GCTAGGGTGTCTGCGCAACTGATCCTCGCGGAGCGTCCTTTCTACGTCCCGTCGGC 1439
Db      752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 772
Qy      1440 GCTGAATCCCGCGGACGACCCGCTCTCGGGCGGCTTGGGGCTCTACCGTCCCTCTTCA 1499
Db      772 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyIlePheArgProLeuLeuAr 792
Qy      1500 TCTGCGCTTCCGGCGCGACACCGGCGCACCTCTCTTTACCGCGTCTCCCGTATGTGCC 1559
Db      792 gLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValPr 812
Qy      1560 TTCTCATCTGCGGACCGTGTGCACTTCGCTTTCACCTCTGACGTCGGATGGAGACCAAC 1619
Db      812 oSerHisLeuProValArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 832
Qy      1620 G 1620
Db      832 o 832

```

Search completed: December 1, 2005, 23:35:15
Job time : 520.5 secs

THIS PAGE BLANK (USPTO)


```
QY 184 ACCCTGCTCGTGTACAGCGGGGTTTTCTCGTTGACAAAGATCCTCAATACGGCA 243
Db 365 ThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrThr 384
QY 244 GAGTCTAGACTCG-GTGGACTTCTCTCAATTTCTAGGGGAGACCCACGCTGTTCCCTG 302
Db 385 GluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-SerTr 404
QY 303 GCCAAATTCGCAAGTCCCAACCTCCAAATCACTCACCAACTCTGTCTCTCCAAATTTGTC 362
Db 404 pProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSe 424
QY 363 CTGGCTATCGCTGAGTGTCTCGGGGTTTTATCATATTCCTCTTCATCTCTGCTGAT 422
Db 424 rTrpLeuSerLeuAspValSerAlaAlaPheTyHisIleProLeuHisProAlaAlaMe 444
QY 423 GCCTCATCTTCTGTGTCTCTGACTACCAAGGTATGTTGCCGTTTGTGCTCTAC 482
Db 444 tProHisLeuLeuValGlySerGlyLeuProA-rgTyValAlaArgLeuSerSerTh 464
QY 483 TTCAGGAACATCAACACGACGCGGCCCATGCAGACCTGCACGACTCTCTGCTCAAG 542
Db 464 rSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSerAr 484
QY 543 GAAACTCTAGCTTCCCTCTGTGTCTGTACAAAACCTTCGGACGGAACTGCACTTGTA 602
Db 484 gAsnLeuTyValSerLeuLeuLeuLeuTyLysThrPheGlyArgLysLeuHisLeuTy 504
QY 603 TTCCCATCCCATCATCTCTGGCTTCGCAAGATTCCTATGGAGAGTGGGCTCAGTCCGTT 662
Db 504 rSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProH 524
QY 663 TCTCTCGGCTCAGTTACTAGTCCCATTTCTCAGTGGTTCTGATGGGCTTCCCCCACTG 722
Db 524 eLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHisCy 544
QY 723 TTTGGCTTTCAGTTATATGATGATGTGTATTTGGGGGCGAATCTGTCAACATCTTGA 782
Db 544 sLeuAlaPheSerTyMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGl 564
QY 783 GTCCCTTTTACCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCTTAA 842
Db 564 uSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAs 584
QY 843 TAAACCAACAGTTGGGGCTACTCCCTTAACCTTCATGGGATATGTAATTTGGAAGTTGGGG 902
Db 584 nLysThrLysArgTrpGlyTySerSerLeuAsnPheMetGlyTyValIleGlySerTrpGl 604
QY 903 TACTTTACCGCAGAACATATTGTACTAAACTCAAGCAATGTTTTCGAAAACCTGCTGT 962
Db 604 yThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuProVa 624
QY 963 AAATAGACCTATTGATTGGAAATGATGTCAAGAATTGGGCTTTTGGGCTTTCCTGC 1022
Db 624 lAsnSerProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAl 644
QY 1023 CCCTTTTACAAATATGGGCTATCCTGCTTGATGCTCTTATATGTCATGTATACAATCTAA 1082
Db 644 aProPheThrGlnCysGlyTyTrpAlaLeuMetProLeuTyAlaCysIleGlnSerLy 664
QY 1083 GCAGGCTTTCATCTTCTCGCCAACTTACAAAGGCTTTCTGTGTAAACAATATCTGAACCT 1142
Db 664 sGlnAlaPheThrPheSerProThrTyLysAlaPheLeuCysLysGlnTyLysLeuAsnLe 684
QY 1143 TTACCGGTTGCGCGCAACGCTCCGCTCTCTCCCAAGTGTGTTGTCGACGCAACCCAC 1202
Db 684 uTyTrpProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProTh 704
QY 1203 TGGATGGGCTTGGCCATAGGCCATCAGCGCATGGGCTGGAAACCTTTCTGCTCTCTGCTGCC 1262
Db 704 rGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLeuPr 724
```

```
QY 1263 GATCCATACTGCGGACTCCTAGCAGCTGTGTTTCTGCTCGACGCGCTCTGGAGCAAAACT 1322
Db 724 oIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLysLe 744
QY 1323 TATCGGAACCGCAACTCTGTTGTCTCTCTCGGAAATACACTCCTCTTCCATGGCTGCT 1382
Db 744 uIleGlyThrAspAsnSerValValLeuSerArgLysTyThrSerPheProTrpLeuLe 764
QY 1383 AGGGTGTGTGCAACTGGATCCTGCGGGGAGCTGCTTTGTCTAGTCCGTCGCGCT 1442
Db 764 uGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyValProSerAlaLe 784
QY 1443 GAATCCCGGGACGACCCGCTCTCGGGCGCTTGGGGCTCTACCGTCCCTCTTCATCT 1502
Db 784 uAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyArgProLeuLeuLe 804
QY 1503 GCGGTTCCGCGCCAGCACGCGGCGCACTCTCTTACGCGGCTCTCCCGCTATGTGCTTC 1562
Db 804 uProPheArgProThrThrGlyArgThrSerLeuTyAlaValSerProSerValProSe 824
QY 1563 TCATCTGCGGACCGGTGTCATCTCGCTTCACCTCTGACGTCGACATGAGACCAACG 1620
Db 824 rHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 843
RESULT 2
S35527
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adr)
C:Species: hepatitis B virus, HBV
A:Variety: subtype adr
C:Date: 09-Dec-1993 #sequence_revision 01-Nov-1996 #text_change 09-Jul-2004
C:Accession: S35527
R:Mukalide, M.; Kumazawa, T.; Hoshi, A.; Kawaguchi, R.; Hikiji, K.
Nucleic Acids Res. 20, 6105, 1992
A:Title: The complete nucleotide sequence of hepatitis B virus, subtype adr (SRADR) and t
A:Reference number: S35527; MUID:93096607; PMID:1461746
A:Accession: S35527
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-843 <NUK>
A:Cross-references: UNIPROT:081107; UNIPARC:UPI000000BED31; EMBL:D12980; NID:g221500; PID:
A:Experimental source: subtype adr
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1992
C:Genetics: 345/3
A:introns:
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase
Alignment Scores:
Pred. No.: 2,44e-198 Length: 843
Score: 2789.00 Matches: 523
Percent Similarity: 98.15% Conservative: 8
Best Local Similarity: 96.67% Mismatches: 9
Query Match: 47.22% Indels: 2
DB: 2 Gaps: 0
US-10-761-006A-1 (1-3215) x S35527 (1-843)
```

```
QY 1 CTCCCAACATTCACCAAGCTCTGCTAGATCCAGGGTGAGGGCCTATATTTCTCTGC 60
Db 304 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 323
QY 61 TGGTGGCTCCAGTTCGGAAACAGTAAACCTGTTCCGACTACTGCTCTCCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyCysLeuThrHisIleVal 343
QY 121 AATCTTCTCGAGCATGGGGACCTGCACCGAACATGGAGAACACAACTCAGGATTCCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGGACCCCTGCTCGTGTATACGCGGGGTTTTTCTCGTTGACAAGAAATCCTCACATAACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
```

```
QY 241 GCAGAGTCTAGACTCTG-GTGGACTCTCTCAATTTCTAGGGGAGCACCACGCTGTC 299
Db |||||
QY 384 ThrGluSerArgLeuValValAlaPheSerGlnPheSerArgGlySerThrHisVal-Se 403
Db |||||
QY 300 CTGGCCAAAATTCGACGTCCCACTCCCAATCACTCACCACCTCTGTCTCCCAATTT 359
Db |||||
QY 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
Db |||||
QY 360 GTCTGCTATCGCTGATGCTGGCGGCTTTTATCATATTCCTCTCATCTCTGCTGC 419
Db |||||
QY 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
Db |||||
QY 420 TATCCCTCATCTCTGTGTGCTCTCTGGACTACCAAGGTATCTGCCGCTTTGCTCCT 479
Db |||||
QY 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaargLeuSerSe 463
Db |||||
QY 480 TACTTCCAGGAACATCAACACACGACGCGGCCCATGCAAGCTGCAACACTCTCTGCTC 539
Db |||||
QY 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
Db |||||
QY 540 AAGGAAACTCTACGTTCCCTCTCTGTTGCTGTACAAAACCTTCGACGGAACTGCACCTT 599
Db |||||
QY 483 rArgAsnLeuTyrValSerLeuLeuValTyrLysThrPheGlyArgLysLeuHisLe 503
Db |||||
QY 600 GTATTCCTCCATCCCATCTCTGGGCTTTTCGCAAGATTCTCTATGGAGTGGGCTCAGTCC 659
Db |||||
QY 503 uTyrSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Db |||||
QY 660 GTTCTCTCTGCTCAGTTTACTAGTGCATTTGTTTCAGTGTCTCTAGGGCTTTCCCCCA 719
Db |||||
QY 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValargAlaPheProHis 543
Db |||||
QY 720 CTGTTTGGCTTTTCAAGTATATGAGTATGAGTGTGGTGTGGGCGCAAGCTGTACAACTCT 779
Db |||||
QY 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
Db |||||
QY 780 TGAGTCCCTTTTACCTCTATATACCAATTTCTTTGCTTTGGGTATACATTTAAACCC 839
Db |||||
QY 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
Db |||||
QY 840 TAATAAAACCAAACTTGGGCTCTCTCCCTTAACTTCAATGGATATGTAATTTGGAAGTTG 899
Db |||||
QY 593 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
Db |||||
QY 900 GGGTACTTTACCGCAGGAACATATTGTACTAAACTCAAGCAATGTTTTGGAACCTGCC 959
Db |||||
QY 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
Db |||||
QY 960 TGTAAATAGACTATTGATTGGAAGTATGTCAAAGAAATTTGGGTCTTTTGGGCTTTCG 1019
Db |||||
QY 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
Db |||||
QY 1020 TGCCCTTTTACACAACTGCTATCCTGCTGCTGATGCTTTATATGATGATATATCAATC 1079
Db |||||
QY 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
Db |||||
QY 1080 TAAAGCAGGCTTTCACTTTCTCGCCAACTTACAAGGCTTTCTGTGTAAACAATATCTGAA 1139
Db |||||
QY 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysGlnGlnTyrLeuHi 683
Db |||||
QY 1140 CTTTACCCGCTGTCGCGCAACCGTCTCTCTGCAAGTGTGTGTCGACGCAACCC 1199
Db |||||
QY 683 sLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheGlyAspAlaThrPr 703
Db |||||
QY 1200 CACTGGATGGGCTTGGCCATAGGCATCAGCGCATGGTGGAACTTTCTGGCTCTCTCT 1259
Db |||||
QY 703 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 723
Db |||||
QY 1260 GCCGATCCATACCTGGGAACTCTAGCAGCTTTGTTTTGCTCGACGCGGTCTGGAGCAAA 1319
Db |||||
QY 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaL 743
Db |||||
QY 1320 ACTTATCGGAACCGCAACTCTGTGTGCTCTCTCGGAAATACACCTCTTCCATGGCT 1379
Db |||||
```

```
Db |||||GlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
QY 1380 GCTAGGCTGTCTGCCCAACTTGGATCTTGGCGGAGCGTCTCTTGTCTACGTCCTCGTCGC 1439
Db |||||
QY 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
Db |||||
QY 1440 GCTGAATCCCGCGAGACCCGCTCTCGGGCCGTTTGGGGCTCTACCGTCCCTTCTTCA 1499
Db |||||
QY 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
Db |||||
QY 1500 TCTGCGTTCGCGGCGACACGCGGCGACCTCTCTTACGCGGTCTCCCGATATNGCC 1559
Db |||||
QY 803 sLeuProPheGlnProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
Db |||||
QY 1560 TTTCTATCTCCCGACCGTGTGCTCTGCTCTCACCTCTCACGTCGTCATGGAGACCAACC 1619
Db |||||
QY 823 oSerHisLeuProValArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
Db |||||
QY 1620 G 1620
Db |||||
QY 843 O 843
Db |||||
RESULT 3
S43491
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adr)
C;Species: hepatitis B virus, HBV
A;Variety: subtype adr
C;Date: 07-Sep-1994 #sequence_revision 22-Nov-1996 #text_change 09-Jul-2004
C;Accession: S43491
R;Lancarevic, I.F.; Zentgraf, H.; Schroeder, C.H.
Nucleic Acids Res. 18, 4940, 1990
A;Title: Sequence of a replication competent hepatitis B virus genome with a prex open te
A;Reference number: S12598; MUID:90370503; PMID:2355664
A;Accession: S43491
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-843 <LON>
A;Cross-references: UNIPROT:Q67952; UNIPARC:UPI00000EE7B3; EMBL:X52939; NID:G457780; PDB:
A;Experimental source: subtype adr
C;Genetics:
A;Gene: P; pol
C;Superfamily: hepatitis virus DNA-directed DNA polymerase
C;Keywords: DNA biosynthesis; nucleotidyltransferase
Alignment Scores:
Pred. No.: 1.89e-195 Length: 843
Score: 2750.00 Matches: 517
Percent Similarity: 97.04% Conservative: 8
Best Local Similarity: 95.56% Mismatches: 15
Query Match: 46.55% Indels: 2
DB: 1 Gaps: 0
US-10-761-006A-1 (1-3215) x S43491 (1-843)
QY 1 CTCCACACATTCACCAAGCTCTAGATCCAGGCTAGGCGCTATATTTTCCTGC 60
Db |||||
QY 304 LeuHisAsnIleProProSerSerAlaArgSerLysSerGluGlyProLeuPheProCys 323
Db |||||
QY 61 TGGTGGCTTCCAGTTCGCGAAACAGTAAACCCCTGTTCCGACTACTGCTCTCCCATATCGTC 120
Db |||||
QY 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
Db |||||
QY 121 AATCTTCTCGAGGACTGGGACCCCTGCACCGAATGAGAACACAAATCAGGATTCCT 180
Db |||||
QY 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
Db |||||
QY 181 AGGACCCCTGCTGTTTACAGCGGGGTTTTCTCGTTGACAAATCCTCACAATACC 240
Db |||||
QY 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisThr 383
Db |||||
QY 241 GCAGAGTCTAGACTCTG-GTGGACTCTCTCAATTTCTAGGGGAGCACCACGCTGTC 299
Db |||||
```

Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnThrGlnVal-Se 403
 QY 300 CTGGCCAAATTCGAGTCCCGAACCTCCATCATTACCAACCTCTGTCTCCAAATTT 359
 Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
 QY 360 GTCCCTGGCTATCGCTGGATGTGTCGCGGCTTTTATCATATTTCCCTTCATCCCTGCTCC 419
 Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
 QY 420 TATGCTCATCTCTTTGTTGTTCTTTCTGGACTACCAAGGTATGTTGGCCGTTGTGCTTC 479
 Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuGlnArgTyrValAlaArgLeuSerSe 463
 QY 480 TACTTCCAGGAACATCAACACACGACGCGGCCATGCAAGACTCGACACTCTCTGCTC 539
 Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
 QY 540 AAGGAAACTCTACGTTTCCCTCTGTGTGCTGTACAAAACCTTCGGACGGAACTGCACATT 599
 Db 483 rLysHisLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisIle 503
 QY 600 GTATTCCCATCCCATCATCTCGGCTTTCCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
 Db 503 uTyrSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
 QY 660 GTTCTCTGGCTCAGTTACTAGTGCATTTGTTTCAGTGGTTCGTAGGCTTTCCCCCA 719
 Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValCysArgAlaPheProHis 543
 QY 720 CTGTTTGGCTTCAGTATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779
 Db 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisIle 563
 QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTGTTCTTTGGGTATACATTAAACCC 839
 Db 563 uGluSerLeuPheThrSerIleThrAsnPheMetLeuSerLeuGlyIleHisLeuAsnPr 583
 QY 840 TAATAAACAACAGTTGGGGCTACTCCCTTAACCTTCATGGGATATGTAATTGGAAGTTG 899
 Db 583 oAsnLysThrLysArgTyrGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
 QY 900 GGGTACTTTACCGCAGACATATTTGCTAAACTCAAGCATGTTTTCGAAACTGCC 959
 Db 603 pGlyThrLeuProGlnGluHisIleValGlnLysLeuLysGlnCysPheArgLysLeuPr 623
 QY 960 TGTAATAGACCTATTGATTGGAAGTATGTCAAAGAATTGGGGTCTTTTGGGCTTTGC 1019
 Db 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
 QY 1020 TGCCCTTTTACAAATGTGGCTATCCTGCCTTGATGCTCTTATATGCAATGATACAATC 1079
 Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
 QY 1080 TAAGAGGCTTTACCTTTCTCGCAACTTACAGGCTTTCTGTGTAAACAATATCTGAA 1139
 Db 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 683
 QY 1140 CTTTTACCGGTTGCGCGCAACGGTCCGCTCTCTGCCAAGTGTTCGAGCGCAACCCC 1199
 Db 683 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
 QY 1200 CACTGGATGGGCTTGGCCATAGGCCATCAGCGCATGGCTGGAACTTTCTGGCTCTCTCT 1259
 Db 703 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 723
 QY 1260 GCGCATCCATCTACGCGAACTCTACGAGCTTGTGTTGCTCGCAGCGGTCTGGAGCAA 1319
 Db 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaTh 743
 QY 1320 ACTTATCGGAACGCAACTCTGTCTCTCTCGGAAATACACCTCTCTTCCATGGCT 1379
 Db 743 rLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProItrPue 763

QY 1380 GCTAGGTGTGCTGCCAACTGGATCCTCGCGGGAGCTCTTTGTCTAGTCCCGTCGCG 1439
 Db 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
 QY 1440 GCTGAATCCCGCGAGCAGACCCGCTCTCGGGGCGGTTTGGGGCTCTACCGTCCCGCTTCTTCA 1499
 Db 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuAr 803
 QY 1500 TCTGCGTTCGCGCGACACACGCGCGCACCTCTCTTTACGCGGTCTCCCGTATGTGCC 1559
 Db 803 gLeuSerPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
 QY 1560 TTTCTCATCTCGCGGACCGTGTGCATCTCGCTTCACCTCTGCGACGTCGATGAGACCACC 1619
 Db 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
 QY 1620 G 1620
 Db 843 o 843
 RESULT 4
 JDVLVS
 DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adr, mutant)
 C:Species: hepatitis B virus, HBV
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
 C:Accession: S04568
 R:Rho, H.M.; Kim, K.; Hyun, S.W.; Kim, Y.S.
 Nucleic Acids Res. 17, 2124, 1989
 A:Title: The nucleotide sequence and reading frames of a mutant hepatitis B virus subtype
 A:Reference number: S04568; MUID:89183619; PMID:2928116
 A:Accession: S04568
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-842 <RHO>
 A:Cross-references: UNIPROT:P31870; UNIPARC:UPI0001297FB; EMBL:X14193
 C:Superfamily: hepatitis virus DNA-directed DNA polymerase
 C:Keywords: DNA biosynthesis; nucleotidyltransferase
 Alignment Scores:
 Pred. No.: 2,65e-194 Length: 842
 Score: 2734.50 Matches: 517
 Percent Similarity: 96.86% Conservative: 7
 Best Local Similarity: 95.56% Mismatches: 15
 Query Match: 46.29% Indels: 3
 DB: 1 Gaps: 1
 US-10-761-006A-1 (1-3215) x JDVLVS (1-842)
 QY 1 CTCGCAACATTCACACAGCTCTCTAGATCCAGGGTGAGGGCGCTATATTTCTCTGC 60
 Db 304 LeuHisHisIleSerProSerProAlaArgSerGlnSerGlnGlyProIlePheSerSer 323
 QY 61 TCGTGGCTCCAGTTCGGGACAGTAACCTCTGCTCCGACTACTGCTCTCCCATATCGTC 120
 Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
 QY 121 AATCTTCTCGAGACTTGGGGACCTCTGCACCGCAACATGGAGAACACAAATCAGGATTCT 180
 Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
 QY 181 AGGACCCCTGCTGTTTACAGCGGGGTTTTCTCGTTGACAAGATCCTCACATAATACC 240
 Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
 QY 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTCTTAGGGGAGGACCCAGCTGTTC 299
 Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
 QY 300 CTGGGCAAAATTCGAGTCCCGCAACCTCCAATCACTCAACACCTCTGTCTCTCCAATTT 359
 Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423

```

QY 360 GTCTGCTATCGTGTGATGCTGCGGCGTTTATCATATTCCTCTTCATCTCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl 443
QY 420 TATCCCTCATCTCTCTGCTGCTCTCTGAGTACCAAGGTATGTTCCCGTGTGCTCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaargLeuSerSe 463
QY 480 TACTTCAGGAACATCAACACACAGCAGCGGGGCCATGCAAGACCTGCAGACCTCTCTC 539
Db 463 rThrSerArgAsnIleAsnHisGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
QY 540 AAGGAAACTCTACGTTTCCCTCTGTTGCTGTCTACAAACCTTCGGACGGAAACTGCACCT 599
Db 483 rArgAsnLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
QY 600 GTATTTCATCCCATCATCTCGGCTTTTCGCAAGATTCCTATGGAGTGGCGCTTCAGTCC 659
Db 503 uTyrSerHisProIleIleLeuGlyPheArgLysLeuProMetGlyGlyLeuSerPr 523
QY 660 GTTTCCTCTGCTCAGTTACTAGTCATTTGTTGCTAGTGTTCGTAGGCTTTCCGCCA 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
QY 720 CTGTTTGGCTTTCAGTTATATGATGATGATGATGATGATGATGATGATGATGATGAT 779
Db 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTCTTTGCTTTGCTGATATACATTTAAACCC 839
Db 563 uGlnSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
QY 840 TAATAAAACCAACCTGGGCTACTCCCTTAACCTCAAGCAATGTTTTCGAAAGTTG 899
Db 583 oAsnLysThrLysArgTyrGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
QY 900 GGGTACTTTACCGAGNACATATGTACTAACTCAAGCAATGTTTTCGAAAGTTG 959
Db 603 pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAATAGACTTATGATGGAAGTATGTCAAGAAATGTTGCTTTGGGCTTTG 1019
Db 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGCCCCCTTTACAAATGTGCTATCTGCTGCTGATGCTTTATGATGATGATGATGATGAT 1079
Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
QY 1080 TAAGCAGCTTTTCACTTTCTCGCNACTTACAGGCTTTCTGTGTAAACATATCTGAA 1139
Db 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi 683
QY 1140 CTTTACCCTGTTGCGCGCAACCGTCCGCTCTCTGCCAAGTGTGTTGTCGACGCAACCCC 1199
Db 683 sLeuTyrProValAlaArg--ArgThrAlaLeuCysGlnValPheAlaAspAlaThrPr 702
QY 1200 CACTGATGGGCTTGGCCATAGGCCATCAGGCCATGCTGGAACCTTTCTGCTCTCTCT 1259
Db 702 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 722
QY 1260 GCCGATCATACTGCGGAACCTCTAGCAGCTTTGTTGCTCGCAGCGGCTGGAGCAA 1319
Db 722 uProIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLy 742
QY 1320 ACTTATCGAACCGCAACTCTGTTGCTCTCTCGGAAATACACCTCTTTCATGCTGCT 1379
Db 742 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 762
QY 1380 GCTAGGCTGTGTCGAACCTGGATCCTGCGGGGACGCTTTGTGTCTACGTCCTCGGC 1439
Db 762 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrTyrPheValTyrValProSerAl 782
QY 1440 GCTGAATCCCGCGGAGGACCGCTCTCGGGGCGGTTTGGGCTCTACCGCTCCCTCTCTCA 1499

```

```

Db 782 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuIleArgProLeuLeuHi 802
QY 1500 TCTGCGCTTCGGCCACACCGGGGCCACCTCTCTTTAGCGGCTCTCCCGTATGTGCC 1559
Db 802 sLeuArgPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 822
QY 1560 TTTCTATCTGCGGACCGTGTGCTTGGCTTCCCTCTCAGCTGTCATGAGAGACACACC 1619
Db 822 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 842
QY 1620 G 1620
Db 842 O 842

RESULT 5
T13473
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (isolate 29Y11HCC)
C:Species: hepatitis B virus, HBV
A:Variety: isolate 29Y11HCC
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: T13473
R:Takahashi, K.; Akahane, Y.; Hino, K.; Ohta, Y.; Mishiro, S.
Arch. Virol. 143, 2313-2326, 1998
A:Title: Hepatitis B virus genomic sequence in the circulation of hepatocellular carcinoma
A:Reference number: Z17684; MUID:99129050; PMID:9930189
A:Accession: T13473
A:Status: preliminary; translated from GB/EMBL/DBDJB
A:Molecule type: DNA
A:Residues: 1-825 (TAK)
A:Cross-references: UNIPROT:Q9YZS3; UNIPARC:UPI00000F4BF5; EMBL:AB014388; NID:93582381;
A:Experimental source: Japanese patient with hepatocellular carcinoma isolate 29Y11HCC
C:Genetics:
A:Gene: P
A:Introns: 303/3
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:
Pred. No.: 7,33e-191 Length: 825
Score: 2688.00 Matches: 504
Percent Similarity: 95.92% Conservative: 13
Best Local Similarity: 93.51% Mismatches: 21
Query Match: 45.51% Indels: 2
DB: Gaps: 0

US-10-761-006A-1 (1-3215) x T13473 (1-825)
QY 7 AACATTCCACCAAGCTCTGCTAGATCCAGGGTCCAGGGGCTATATTTTCTGCTGTGG 66
Db 288 HisLeuSerThrThrLysArgGlnSerSerSerGlyHisAlaValGluThrCysTrpTrp 307
QY 67 CTCAGTTCGGAAACAGTAAACCTGTTCGAGTACTGCTCTCCATATATGTCATCTT 126
Db 308 LeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAsnLeu 327
QY 127 CTCAGGACTGGGACCGCTCCGACACATGAGACACACATCAGGATTCCTAGGACC 186
Db 328 LeuGlnAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArgThr 347
QY 187 CTCTCTGTTTACAGCGGGGTTTTCTCTGTTGACCAAGATTCCTCACAATACCGCAGAG 246
Db 348 ProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThrThrGlu 367
QY 247 TCTAGACTCTG-GTGGACTCTCTCAATTTCTAAGGGGAGCACCACGTTGCTGCGCC 305
Db 368 Ser**LeuValValAspPheSerGlnPheSerArgGlySerThrGlnVal-SerTrpPr 387
QY 306 AAAATTGCGAGTCCCAACCTCCATCTACCTACCAACCTCTTGCTCCTCCAAATTTGCTCG 365
Db 387 oLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSerTr 407
QY 366 GCTATCGCTGATGTGTCTGCGGGCGTTTTTATCATATTTCTCTTCATCTGCTGCTATGCC 425

```


407	Db	pLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAlaMetPr	427
426	QY	TCATCTCTTTGTTGGTTCTTCTCGAGCTACCAAGGTATGTTGCCGTTTGCTCTACTTC	485
427	Db	oHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerThrSe	447
486	QY	CAGGAACATCAACACACAGCAGCGGGCCATGCAAGACCTGCACGACTCTCTGCTCAAGAA	545
447	Db	rArgAsnIleAsn**GlnHisGlyThrMetGlnAspLeuHisAspSerCysSerIysHis	467
546	QY	ACTCTAGTTTCCTCTTGTGTCTGTATACAAAACTTCGGACGGAACACTGCATGTATTC	605
467	Db	sLeuTyrValSerLeuLeuLeuLeuTyrIysThrPheGlyArgIysLeuHisLeuTyrSe	487
606	QY	CCATCCCATCATCTCGGGCTTTCGCAAGATCTCTATGGGAGTGGGCTCAGTCCGTTCT	665
487	Db	rHisProIle**LeuGlyPheArgIysIleProMetGlyValGlyLeuSerProPheLe	507
666	QY	CTGGGCTCAGTTACTAGTCGCCATTTGTTTCAGTGGTTTCGTAGGCTTTCCTCCCACTGTTT	725
507	Db	uLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHisCysLe	527
726	QY	GGCTTTCAGTTATATGATGATGTGGTATTTGGGGCGAAGCTCTGTACAACTCTTGATC	785
527	Db	uAlaPheSerTyrMetAspAspValValLeuGlyAlaIysSerValGlnHisLeuGluAl	547
786	QY	CTTTTTCCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCTAATAA	845
547	Db	aLeuTyrThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAsnLy	567
846	QY	AACCAACGTTGGGGCTACTCCCTTAACTTCATGGGATATGTAATGGAAAGTTGGGCTAC	905
567	Db	sThrIysArgTyrGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTyrGlyTh	587
906	QY	TTTACCGCAGGAACATATTGTACTTAAACATCAAGCAATGTTTTCGAAAACTGCTGTAAA	965
587	Db	rLeuProGlnAspHisIleValGlnIysLeuIysGlnCysPheArgIysLeuProValAs	607
966	QY	TAGACCTATTGATTGGAAAGTATGTCNAAGAAATGTGGGTCTTTTGGGCTTCTGCCCC	1025
607	Db	nArgProIleAspTyrIysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAlaPr	627
1026	QY	TTTTACACAATGGGCTATCTCGCTTGATGCTCTTATATGATGTATACAACTCAAGCA	1085
627	Db	oPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerIysGI	647
1086	QY	GGCTTTTCACTTTCTCGCAACTTACAGGCCTTTCTGTGTAAACAATATCTGAACCTTTA	1145
647	Db	nAlaPheThrPheSerProThrTyrIysAlaPheLeuCysIysGlnTyrLeuAsnLeuTy	667
1146	QY	CCCCGTTGCCGGAACCGTTCGGTCTCTGCCAAGTGTTCGTGACGCAACCCCCACTGG	1205
667	Db	rProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProThrGI	687
1206	QY	ATGGGGCTTGGCCATAGGCCATCAGCGCATGGCTGGAACTTTCTGCTCCTCTGCCGAT	1265
687	Db	yTyrGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaAlaProLeuProII	707
1266	QY	CCATCTCGCGAACTCTCTAGCAGCTTGTTTTGCTCGCAGCGGTCTGGAGCAAACTTAT	1325
707	Db	eHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaIysLeuII	727
1326	QY	CGAAACCGACAACCTGTGTGTCTCTCTCGGAATAACACCTCTCTTCATGGCTGTAGG	1385
727	Db	eGlyThrAspAsnSerValValLeuSerArgIysTyrThrSerPheProTyrLeuLeuGI	747
1386	QY	GTCTGTGCGCAACTGGAATCTCTCGCGGAGCGTCTTTGTCTAGTCCCGCTCGGCTGAA	1445
747	Db	yCysAlaAlaAsnTrpIleuArgGlyThrSerPheValTyrValProSerAlaLeuAs	767
1446	QY	TCCCGCGGACGACCCGCTCTCGGGGCCGTTTGGGGCTCTACCGTCCCCCTTCTCATCTGCC	1505

767 nProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHisLeuPr 787

1506 GTTCCGGCCGACACCGGGCGCACCTCTCTTTACGCGGTCTCCCGTATGTGCTTCTCA 1565

787 oPheArgProThrThrGlyAArgThrSerLeuTyrAlaValSerProSerValProSerHi 807

1566 TCTCCGGACCGTGTGCACTTCGTTTCACCTCTGCATCGATGGAGACACCG 1620

807 sLeuProAspArgValHisPheAlaSerProLeuHisValAlaTArgProPro 825

RESULT 6

TL3468

DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (isolate 09D09HCC)

C:Species: hepatitis B virus, HBV

A:Variety: isolate 09D09HCC

C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004

C:Accession: TL3468

R:Takahashi, K.; Akahane, Y.; Hino, K.; Ohta, Y.; Mashiro, S.

Arch. Virol. 143, 2313-2326, 1998

A:Title: Hepatitis B virus genomic sequence in the circulation of hepatocellular carcinoma

A:Reference number: Z17684; MUID:99129050; PMID:9930189

A:Accession: TL3468

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-827 >TAK>

A:Cross-references: UNIPROT:Q9Y2U3; UNIPARC:UPI00000F479C; EMBL:AB014368; NID:G3551304; I

A:Experimental source: Japanese patient with hepatocellular carcinoma isolate 09D09HCC

C:Genetics:

A:Gene: P

A:Introns: 303/3

C:Superfamily: hepatitis virus DNA-directed DNA polymerase

C:Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:

Pred. No.:	3,78-190	Length:	827
Score:	2678.50	Matches:	504
Percent Similarity:	95.91%	Conservative:	12
Best Local Similarity:	93.68%	Mismatches:	18
Query Match:	45.34%	Indels:	5
DB:	2	Gaps:	1

US-10-761-006A-1 (1-3215) x TL3468 (1-827)

Qy 19 AGCTCTGTAGATCCAGGGGTGAGGGG-----CCTATATTTTCTCGTGGTGGCTC 69

Db 291 ThrSerLysArgGlnSerSerSerGlyHisAlaValGluLeuHisProCysTrpTrpLeu 310

Qy 70 CAGTTCGGGAACAGTAACCTGTTCGACTACTCGCTCTCCCATATCTCAATCTTCTC 129

Db 311 GlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisileValAsnLeuLeu 330

Qy 130 GAGACTGGGACCTGCACCGACATCGAGAACACACATCAGGATTCCTTAGGACCCCT 199

Db 331 GluAspTrpGlyProCysThrGluHisGlyGluHisAsnileArgileProargTrpPro 350

Qy 190 GCTCGTGTACAGGGGGGTTTTCTCGTTGACAAAGATTCCTCACAATACCGCAGAGTCT 249

Db 351 AlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrThrGluSer 370

Qy 250 AGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGGAGACCCACGCTGTTCCTGGCCAAA 308

Db 371 ArgLeuValValAspPheSerGlnPheSerArgGlySerThr***Val-SerTrpProLy 390

Qy 309 ATTGCGAGTCCCCACCTCCATCACTACCAACCTCTGTCTCCCAATTTGCTCGGCT 368

Db 390 sPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSerTrpLe 410

Qy 369 ATCGCTGATGTCTCGGGGGGTTTTATCATATTCTCTTCATCTCGCTGCTATGCTCA 428

Db 410 uSerLeuAspValSerAlaAlaPheThyHisileProLeuHisProAlaAlaMetProHi 430

Qy 429 TCTTCTGTGTTGTTCTCTGGAGTACCAAGGATGTTGCCCGTTGCTCTCTACTTCCAG 488

Db 430 sLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerThrSerAr 450
Qy 489 GAACATCAACACACAGCAGCGGCCATGCAAGACCTGCAGACTCCCTGCTCAAGAACT 548
Db 450 gAsn***AsnTyrGlnHisGlyThrMetGlnAapLeuHisGlySerCysSerArgAenLe 470
Qy 549 CTACGTTTCCCTCTGTGTGTACAAAACCTTCGGACGGAACCTGCAGCTGATTTCCCA 608
Db 470 uTyrrValSerLeuLeuLeuLeuTyrrPheGlyArgLysLeuHisLeuTyrrSerHi 490
Qy 609 TCCCATCATCTCGGCTTTTCGCAAGATTCCTATCGGAGTGGCTCAGTCGCTTCTCT 668
Db 490 sProlleleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPheLeu 510
Qy 669 GGCTCAGTTACTAGTCGCAATTTGTCAGTGGTTCGTAGGCTTTCCCCCACTGTTGGC 728
Db 510 uAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHisCysLeuAl 530
Qy 729 TTTTCAGTTATATGATGATGGTATTGGGGCGAAGTCTGTACAAATCTTGAGTCCCT 788
Db 530 aPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGluSerLe 550
Qy 789 TTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCCCTAATAAAC 848
Db 550 uPheThrAlaIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAenProHisLysTh 570
Qy 849 CAACAGTTGGGCTACTCCCTTAATCTCATGGATATGTAATTCGAAGTTGGGTACTTT 908
Db 570 rLysArgTrpGlyTyrSerLeuAenPheMetGlyTyrValIleGlySerTrpGlyThrLe 590
Qy 909 ACCCGAGAACATATGTACTAAACCAACCAATGTTTTCGAAACCTGCTGTAAATAG 968
Db 590 uProlGlnHisIleValGlnLysIleLysGlnCysPheArgLysLeuProValAsnAr 610
Qy 969 ACCTATTGATTGGAAGTATGTCAAAGAAATGTGGGTCTTTTGGGCTTTGCTGCCCTTT 1028
Db 610 gProlleasprLysValCysGln**IleValGlyLeuLeuGlyPheAlaAProPh 630
Qy 1029 TACAAATGTGGCTATCTGCTGATGCTGATGCTTATATGCAATGTATACATTAAGCAGGC 1088
Db 630 eThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnAlaLysGlnAl 650
Qy 1089 TTTTCACTTCTCGCCACTACAGGCTTCTGTGTAAACAATATCTGAACCTTTACCC 1148
Db 650 aPheThrPheSerProThrTrpLysAlaPheLeuCysLysGlnTrpLeuHisLeuTyrrPr 670
Qy 1149 GTTGGCCCGCAACGGTCCGCTCTCTGCAAGTGTCTGCTGACGCAACCCCACTGGATG 1208
Db 670 oValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProThrGlyTr 690
Qy 1209 GGGCTTGGCCATAGCCATCAGCGCATGGCTGGAACCTTTCTGGCTCTCTGCGCATCCA 1268
Db 690 pGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLeuProIleHi 710
Qy 1269 TACTCGGAATCTCTAGCAGCTGTGTTGCTCGAGCGGTCTGAGCAAACTTATCGG 1328
Db 710 sThrAlaGluLeuLeuAlaAaCysPheAlaArgSerArgSerGlyAlaLysLeuIleGl 730
Qy 1329 AACCGACAATCTGTGTGCTCTCTCGGAATATACACTCTCTTTCATGGCTGCTAGGGTG 1388
Db 730 yThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLeuLeuGlyCy 750
Qy 1389 TGCTGCCAATCGATCTCGCGGAGCTCTTTGTCGTACGTCGCGTGGCGCTGAATCC 1448
Db 750 sAlaAlaAenTrpIleLeuArgGlyThrSerPheValTyrValProSerAlaLeuAenPr 770
Qy 1449 CCGGACGACCGCTCTCGGGCCGTTTGGGCTCTACCGTCCCTCTTCTCATCTCGCGTT 1508
Db 770 oAlaAspAspProSerArgLysArgLeuGlyLeuTyrrArgProLeuLeuHisLeuProPh 790
Qy 1509 CCGGCGGACCGGCGGCGCACTCTCTTACCGGTCTCCCGTATGTGCCTTCTCATCT 1568
Db 790 eArgProThrThrGlyArgThrSerLeuTyrrAlaValSerProSerValProSerHisLe 810

Qy 1569 GCGGACCGGTGTGCACCTTGCTTACCTTGCACGTGCGATGAGACCCG 1620
Db 810 uProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 827

RESULT 7

S71785
DNA-directed DNA polymerase (BC 2.7.7.7) - hepatitis B virus (subtype ayw, isolate patient
C;Species: hepatitis B virus, HBV
A;Variety: subtype ayw, isolate patient C1005
C;Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 09-Jul-2004
C;Accession: S71785
R;Preiller-Adams, S.; Schlayer, M.J.; Peters, T.; Hettler, F.; Gerok, W.; Rasenack, J.
submitted to the EMBL Data Library, March 1993
A;Description: Identification and sequence analysis of hepatitis B virus DNA in immunolob
A;Reference number: S32202
A;Accession: S71785
A;Molecule type: DNA
A;Residues: 1-832 <PRE>
A;Cross-references: UNIPROT:Q8JN11; UNIPROT:Q9IF40; UNIPROT:Q96846; UNIPROT:Q67892; UNIP
PROT:Q9QAG0; UNIPROT:Q9QAF3; UNIPROT:Q80X01; UNIPROT:O11885; UNIPROT:O56654; UNIPROT:O568
A;Experimental source: subtype ayw, isolate patient C1005
C;Superfamily: hepatitis virus DNA-directed DNA polymerase
C;Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:
Pred. No.: 2,42e-188 Length: 832
Score: 2654.00 Matches: 492
Percent Similarity: 95.56% Conservative: 25
Best Local Similarity: 90.94% Mismatches: 23
Query Match: 44.93% Indels: 2
DB: Gaps: 0

US-10-761-006a-1 (1-3215) x S71785 (1-832)

Qy 1 CTCACAACTTCCACCAAGCTCTGTAGATCCAGGTCAGGGGCTATATTTTCCTGC 60
Db 293 LeuHisAenPheProProAenSerAlaArgSerGlnGlyGluArgProValPheProCys 312
Qy 61 TGTGGCTCCAGTCCGGAACAGTAAACCTGTTCGCACTACTGCTCTCCCATATGCT 120
Db 313 TrpTrpLeuGlnPheArgAenSerLysProCysSerAspTyrCysLeuSerHisIleVal 332
Qy 121 AATCTTCTCAGGACTGGGACCTGCACCGAAATCATGAGAACACAAATCAGGATTCCT 180
Db 333 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisIleArgIlePro 352
Qy 181 AGGACCCCTGCTCGTGTATACAGCGGGTGTTCCTGTGACAAAGATCCTCACAATACC 240
Db 353 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAenProHisAenThr 372
Qy 241 GCAGACTAGACTCTG- GTGGACTTCTCAATTTCTAGGGGAGACCCACGCTGTC 299
Db 373 AlaGluSerArgLeuValAlaAspPheSerGlnPheSerArgLysAenHisArgVal-Se 392
Qy 300 CTGCGCAAAATTCAGTCCCAACCTCCAACTCACTCAACCAACCTTGTCTCCCAATTT 359
Db 392 rTrpProlLysPheAlaValProAenLeuGlnSerLeuThrAsnLeuLeuSerAenLe 412
Qy 360 GTCTCGGCTACTCGTGGATGTCTCGGGGGTGTATATTCATATTCCTCTTCCTGCTGC 419
Db 412 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaSe 432
Qy 420 TATCGCTCATCTTCTGTTGGTCTTCTGAGTACCAAGGTATGTCGCCCTTTGTCCTC 479
Db 432 rMetProHisLeuLeuValGlySerThrGlyLeuSerArgTyrValAlaArgValSerSe 452
Qy 480 TACTTTCAGGAACATCAACACACGAGCGGGCCATGCAAGACCTGCAAGCTCCTCCTGC 539
Db 452 rAsnSerArgIlePheAenHisGlnArgGlyThrMetGlnAenLeuHisAspTyrCysSe 472
Qy 540 AAGGAAACTCTACGTTTCCCTCTTGTGCTGTACAAACCTTCGAGCGGAACTGCACCT 599

Db 472 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrGlnThrPheGlyArgLysLeuHisLe 492
QY 600 GTATTCCCATCCCATCATCTCGGCTTTTCGCAAGATTCCTATGGGAGTGGGCTCAGTCC 659
Db 492 uTyrSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 512
QY 660 GTTTCCTCCTGGCTCAGTTTACGTAGTCCCATTTGTTTCAGTGGTTCGTAGGCTTCCCCCA 719
Db 512 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 532
QY 720 CTGTTTGGCTTTTCAGTTATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779
Db 532 sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe 552
QY 780 TGAGTCCCTTTTACCTCTATTAACCAATTTCTTTTGTCTTTGGGATATACATTTAAACCC 839
Db 552 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 572
QY 840 TAATAAACCAACGTTGGGCTACTCCCTTAACCTCATGGGATATGTAATGGAGTTG 899
Db 572 oAsnLysThrLysArgTrpGlyTyrSerLeuHisPheMetGlyTyrValIleGlySerTy 592
QY 900 GGGTACTTTACCCAGCAACATATCTACTTAAACCTCAAGCAATGTTTTCGAAAACCTGCC 959
Db 592 rGlySerLeuProGlnAspHisIleIleGlnLysIleLysGluCysPheArgLysLeuPr 612
QY 960 TGTAAATAGACCTATTGATTGGAAAGATGATGTCAAAGAAATTGGGGTCTTTGGGCTTTGC 1019
Db 612 oIleAsnArgProIleAsePrLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 632
QY 1020 TGCCCTTTTACCAATGGCTATCTGCTGCTGATGCTCTTATATGATGATGATGATGATGAT 1079
Db 632 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 652
QY 1080 TAAGCAGGCTTTTCACTTTCTCGCAACTTACAAGGCTTTCTGTGTAAACAATATCTGAA 1139
Db 652 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 672
QY 1140 CCTTTACCCCGTTGCCGCAACGGTCCGGTCTCTGCCAAGTGTTCCTGACCAACCC 1199
Db 672 nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 692
QY 1200 CACTGATGGGCTTGGCCATAGCCATCAGCCATGGCTGGAACTTTCTGCTCCTCT 1259
Db 692 oThrGlyTrpGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheLeuAlaProLe 712
QY 1260 GCCGATCCCATCTGCGAACTCCTAGCAGCTTCTTTTTCGCGAGCGGTCTCGAGCAAA 1319
Db 712 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAs 732
QY 1320 ACTTATCGGAACCGACAACCTCTGTTGCTCTCTCGGAATATACCTCTCTTCCATGGCT 1379
Db 732 nIleLeuGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 752
QY 1380 GCTAGGCTGCTGCCAATGGATCTCGCGGAGCTCCTTTGCTAGTCCCGTGGCC 1439
Db 752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 772
QY 1440 GCTGAATCCCGGACGACCGCTCGGGCCGTTTGGGGCTCTACCGTCCCTCTCTCA 1499
Db 772 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyIlePheArgProLeuLeuAr 792
QY 1500 TCTGCCGTTCCGCGCAACCGGGCGACCTCTCTTTTACCGGTCTCCCGGTATGTGCC 1559
Db 792 gLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValPr 812
QY 1560 TTCTCATCTCCCGACCGTGTGCACTTCGCTTCACTCTGACGTGCGATGGAGACACC 1619
Db 812 oSerHisLeuProValArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 832
QY 1620 G 1620
Db 832 o 832

RESULT 8

JDLVLVB
DNA-directed DNA polymerase (BC 2.7.7.7) - hepatitis B virus (subtype ayw, strain PHB320)
C:Species: hepatitis B virus, HBV
A:Note: host Homo sapiens (man)
C:Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004
C:Accession: A00703
R:Bichko, V.; Pushko, P.; Dreilina, D.; Pumpen, P.; Gren, E.
FEBS Lett. 185, 208-212, 1985
A:Title: Subtype ayw variant of hepatitis B virus: DNA primary structure analysis.
A:Reference number: A05237; MUID:85204397; PMID:3996597
A:Accession: A00703
A:Molecule type: DNA
A:Residues: 1-832 <BI>
A:Cross-references: UNIPROT:P03156; UNIPARC:UPI0000171088; GB:X02496; NID:962280; PIDN:CU
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:
Pred. No.: 3.4e-188 Length: 832
Score: 2652.00 Matches: 494
Percent Similarity: 95.38% Conservative: 22
Best Local Similarity: 91.31% Mismatches: 24
Query Match: 44.90% Indels: 2
DB: 1 Gaps: 0

US-10-761-006a-1 (1-3215) x JDLVLVB (1-832)

QY 1 CTCACCAACATTCCACCAAGCTCTGTAGATCCAGGGTGAGGGCTATATTTTCTCTGC 60
Db 293 LeuHisAsnLeuProAsnSerAlaArgSerGlnSerGluArgProValPheProCys 312
QY 61 TGGTGGCTCCAGTTCGGGAACAGTAAACCTGTTCCGACTACTGCTCTCCCATATCGTC 120
Db 313 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal 332
QY 121 AATCTCTCAGAGCTGGGAGCCCTGCACCAACATGAGGAACACACACATCAGGATCCCT 180
Db 333 AsnLeuLeuGluAspTrpGlyProCysAlaGluHisGlyGluHisIleArgIlePro 352
QY 181 AGACCCCTCTGCTGTGTACAGCGGGGTTTTCTGTTGACAAAGAAATCCTCACATACC 240
Db 353 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 372
QY 241 GCAGAGCTTAGACTCTG-GTGACTTCTCTCAATTTCTAGGGGAGAGCCACCGTGTTC 299
Db 373 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnTyrArgVal-Se 392
QY 300 CTGGCCAAATTCGCAGTCCCAACCTCCAATCAGCTCCCAACCTCTCTCTCCAATTT 359
Db 392 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerAsnLe 412
QY 360 GTCTGGCTATCGCTGAGTGTCTGCGCGGTTTTATCATATTTCTCTTTCATCCTCTCTGC 419
Db 412 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl 432
QY 420 TATGCTCATCTCTTGTGTTGTTCTTCTGAGTACCAAGGTATGTTGCCGTTGTCTTC 479
Db 432 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 452
QY 480 TACTTCCAGAACATCAACCAACGACGCGGCCATGCAAGACCTTCGAGCACTCCTGCTTC 539
Db 452 rAsnSerArgIlePheAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 472
QY 540 AAGGAAACTCTAGTTTCCCTCTTGTGCTGTACAAAACCTTCGAGCGAAACCTGCACATT 599
Db 472 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrGlnThrPheGlyArgLysLeuHisLe 492
QY 600 GTATTCCCATCCCATCATCTCGGCTTTCGCAAGATTCCTATGGGAGTGGGCTCAGTCC 659
Db 492 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 512

```
Qy 660 GTTCTCTCGCTCAGTTTACTAGTGCATTTGTTTCTAGTGTCTGTTAGGCGCTTTCCGCCA 719
Db 512 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHI 532
Qy 720 CTGTTTCTGCTTTTCACTTATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779
Db 532 sCyLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 552
Qy 780 TGAGTCTCTTTTACCTCTATTACCAATTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 839
Db 552 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 572
Qy 840 TATAAACAACAAGCTGGGGCTACTCCCTTAACCTCATGCGATATGTAATTTGGAAGTTG 899
Db 572 oAsnLysThrLysArgTrpGlyTyrSerLeuHisPheMetGlyTyrValIleGlyCysTy 592
Qy 900 GGGTACTTTTACCGCAGGAACATATTGTACTAAAACTCAAGCAATGTTTTCGAAAACCTGCC 959
Db 592 rGlySerLeuProGlnAspHisIleIleGlnLysIleLysGluCysPheArgLysLeuPr 612
Qy 960 TGTAAATAGACCTATTGATTGGAAGATATGTCAAGAATTGTGGGTCTTTTGGCTTTGCC 1019
Db 612 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 632
Qy 1020 TGCCCTTTTACAAATGTGGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079
Db 632 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 652
Qy 1080 TAACGAGCTTTTCACTTTCTCGCAACTTACAAGGCTTTCTGTGTAAACAATATCTGAA 1139
Db 652 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 672
Qy 1140 CTTTATCCCGTTCGCCGCAACCGTTCGGTCTCTGCAAGTGTGTGCTGACGCAACCC 1199
Db 672 nLeuTyrProValAlaArgGlnAArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 692
Qy 1200 CACTGGATGGGCTTGGCCATAGCCATACGCGATGCTGGAACCTTTCTGGCTCTCTCT 1259
Db 692 oThrGlyTrpGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheLeuAlaArgLe 712
Qy 1260 GCCATCCATCTCGGAACCTCTAGCAGCTTGTGTTGCTGCGACCGGCTCGGAGCAAA 1319
Db 712 uProIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaAs 732
Qy 1320 ACTTATCGGAACGACAACTCTGTGCTCTCTCGGAAATACACCTCTTTCATGCT 1379
Db 732 nIleLeuGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerTyrProTrpLe 752
Qy 1380 GCTAGGGTGTCTGCACTGATCTGCGGAGCGTCTTGTGCTAGCTCCGCTCGGC 1439
Db 752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 772
Qy 1440 GCTGAATCCCGCAGCAGCCGCTCTCGGGGCGGTTTGGGGCTCTACCGTCCCTTCTTCA 1499
Db 772 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuSerArgProLeuLeuAr 792
Qy 1500 TCTGCGGTTCGGCGCAGCAGCGGGCGCACTCTCTTTTACGGGGTCTCCCGGTATGTC 1559
Db 792 gLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValPr 812
Qy 1560 TTCTCATCTGCGGACCGTGTGCATCTGCTTCACTTCACTCTGACGTCGCGATCGGACCA 1619
Db 812 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpAArgProPr 832
Qy 1620 G 1620
Db 832 o 832
```

RESULT 9

S47406

DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype ayw4)

C:Species: hepatitis B virus, HBV

A:Variety: subtype ayw4

C:Date: 23-Nov-1994 #sequence_revision 22-Nov-1996 #text_change 09-Jul-2004

C:Accession: S47406

R:Plucieniczak, A.

submitted to the EMBL Data Library, August 1994

A:Description: Molecular cloning and sequencing of two complete genomes of polish isolat

A:Reference number: S47404

A:Accession: S47406

A:Molecule type: DNA

A:Residues: 1-832 <PLU>

A:Cross-references: UNIPROT:Q67892; UNIPARC:UPI00000EF97E; EMBL:Z35716; NID:G527435; Pibk

A:Experimental source: subtype ayw4

C:Genetics:

A:Gene: P

C:Superfamily: hepatitis virus DNA-directed DNA polymerase

C:Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:

Pred. No.:	4,03e-188	Length:	832
Score:	2651.00	Matches:	494
Percent Similarity:	95.19%	Conservative:	21
Best Local Similarity:	91.31%	Mismatches:	25
Query Match:	44.88%	Indels:	2
DB:	1	Gaps:	0

US-10-761-006A-1 (1-3215) x S47406 (1-832)

```
Qy 1 CTCACAACTTCCACCAGCTCTGTAGATCCAGGGTGGGGCTATATTTCTCTGC 60
Db 293 LeuHisAsnLeuProAsnSerArgSerGlnGlyGluArgProAlaPheProCys 312
Qy 61 TGTGGCTCCAGTTCGGAAACAGTAAACCTGTTCCGACTACTGCTCTCTCCCATATCCTC 120
Db 313 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisLeVal 332
Qy 121 AATCTTCTCAGAGCTGGGACCTGACCGAACATGAGAGAACACACATCAGGATTCCT 180
Db 333 AsnLeuLeuGluAspTrpGlyProCysAlaGluHisGlyGluHisIleArgIlePro 352
Qy 181 AGGACCTCTCGTGTGTACAGCGGGGTTTCTGCTGACAAAGATCTCTCACAATACC 240
Db 353 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 372
Qy 241 GCAGAGTCTAGACTCTG-GTGGACTCTCTCAATTTTCTAGGGGGAGACCCACGTGTC 299
Db 373 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnTyrArgVal-Se 392
Qy 300 CTGGCCAAAATTCGACTCCCAACCTCCAACTCACTCACTCACTCTGCTCTCAATTT 359
Db 392 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 412
Qy 360 GTCTGGCTATCGCTGATGCTGCGGGGTTTATCATATCTCTTCACTCTGCTGC 419
Db 412 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl 432
Qy 420 TATGCTCTATCTTCTGTTGTTCTTCTGGACTACCAAGGTATGTCCTGCTGCTCCTC 479
Db 432 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 452
Qy 480 TACTTCCAGGAACATCAACCAACGACGCGGGCCATGCAAGACCTGCAACGACTCCTGCTC 539
Db 452 rAsnSerArgIlePheAsnAsnGlnArgGlyThrMetGlnAsnLeuHisAspTyrCysSe 472
Qy 540 AAGAAACTCTACCTTTTCTGTTGCTGTACAAAACCTTCGAGCGGAACCTGCACCT 599
Db 472 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrGlnThrPheGlyArgLysLeuHisLe 492
Qy 600 GTATTCCCATCCATCATCTGCGCTTTCGCAAGATTCCTATGGAGTGGGCGCTCAGTCC 659
Db 492 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 512
Qy 660 GTTCTCTCGCTCAGTTTACTAGTGCATTTGTTTCTAGTGTGTTCTAGGCTTTCCGCCA 719
Db 512 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHI 532
```



```
Db 453 nSerArgIleLeuAanAanGlnHisGlyThrMetProAspLeuHisAspTyrCySerAr 473
Qy 543 GAAACTCTAGCTTCCCTCTGTTGCTGCTACAAACCTTCGGACGGAACCTGCACTTGTA 602
Db 473 gAsnLeuTyrValSerLeuLeuLeuTyrGlnThrPheGlyArgLysLeuHisLeuTy 493
Qy 603 TTCCCATCCCATCATCTCGGCTTTCGCAAGATTCTCTATGGAGTGGGCTTCAGTCCGTT 662
Db 493 rSerHisProIleleuGlyPheArgLysIleProMetGlyValGlyLeuSerProPh 513
Qy 663 TCTCCTGGCTCAGTTTACTAGTGCATTGTTCAGTGTTCGTAGGTCTTCCCACTG 722
Db 513 eLeuLeuAlaGlnPheThrSerAlaIleCySerValValArgArgAlaPheProHisCy 533
Qy 723 TTTCGCTTTCAGTTATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 782
Db 533 sLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGl 553
Qy 783 GTCCCTTTTACCTCTATTACCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 842
Db 553 uSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAs 573
Qy 843 TAAACCAAAAGCTGGGCTACTCCCTTAACTTCATGGATATGATTAATGGAAGTTGGG 902
Db 573 nLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlyCysTyrGl 593
Qy 903 TACTTTACCGCAGGACATATGTTACTAAACTCAAGCAATGTTTTCGAAAACCTGCTGT 962
Db 593 ySerLeuProGlnGluHisIleIleGlnLysIleGlyLeuGlyCysPheArgLysLeuProIl 613
Qy 963 AAATAGACTTATCATGGAAGTATGTCAAAGAAATGCTGGCTCTTTGGGCTTGTGTC 1022
Db 613 eAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAl 633
Qy 1023 CCCTTTTACAAATGTGCTATCTGCTTGCCTTGCCTTATATGATGATGATGATGATGAT 1082
Db 633 aProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerLy 653
Qy 1083 GCAGCTTTTCACTTCTCGCAACTTACAGGCTTCTGCTGTAACATATCTGAACCT 1142
Db 653 sGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsnLe 673
Qy 1143 TTATCCCGTTCGCGCAACGTCGCTCTGCAAGTGTTCGCAAGTGTTCGCAAGTGTTCG 1202
Db 673 uTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrProth 693
Qy 1203 TGGATGGGCTTGGCATATGCGCATGCGCATGCGCTGGAACCTTCTGCTGCTCTGCTG 1262
Db 693 rGlyTrpGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheSerAlaProLeuPr 713
Qy 1263 GATCATATCTCGGAACTCTCTAGCAGCTTGTGCTGCGAGCGGTCTGGAGCAAACT 1322
Db 713 oIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaAsnIl 733
Qy 1323 TATCGAACCCGACACTCTGTGCTCTCTCGGAATACACCTCTTTCATGCTGCTCT 1382
Db 733 eIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLeuLe 753
Qy 1383 AGGCTGTGCTGCCAATCTGATCTCGCGGGGACGTCCTTGTCTACGTCGCGTGGCGCT 1442
Db 753 uGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAlaLe 773
Qy 1443 GAATCCCGCGGACGACCGCTCTCGGGCGGTGTTGGGCTCTACCGTCCCTCTTCTCATCT 1502
Db 773 uAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuSerArgProLeuLeuArgLe 793
Qy 1503 GCCGTTCGGCGGACCAAGGGGCGACCTCTCTTACCGGCTCTCCCGTATGTCGCTTC 1562
Db 793 uProPheArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValProSe 813
Qy 1563 TCATCTGCGGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
Db 813 rHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 832
```

RESULT 11

```
JDVLJ3
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adw, strain Indonef
C;Species: hepatitis B virus, HBV
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C;Accession: F28925
R;Okamoto, H.; Tada, F.; Sakugawa, H.; Sastrosowignjo, R.I.; Imai, M.; Miyakawa, Y.; Ma
J. Gen. Virol. 69, 2575-2583, 1988
A;Title: Typing hepatitis B virus by homology in nucleotide sequence: comparison of surfa
A;Reference number: J0253; MUID:89010694; PMID:3171552
A;Accession: F28925
A;Molecule type: DNA
A;Residues: 1-843 <OKA>
A;Cross-references: UNIPROT:P17393; UNIPARC:UPI00001297F8; GB:D00331; NID:g221499
C;Superfamily: hepatitis virus DNA-directed DNA polymerase
C;Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:
Pred. No.: 1-58e-187 Length: 843
Score: 2643.00 Matches: 491
Percent Similarity: 95.38% Conservative: 25
Best Local Similarity: 90.76% Mismatches: 24
Query Match: 44.74% Indels: 2
DB: 1 Gaps: 0

US-10-761-006A-1 (1-3215) x JDVLJ3 (1-843)
Qy 1 CTCACAAATTCACCAACCTCTGCTAGATCCAGGTCAGGGGCTATATTTTCCTGC 60
Db 304 LeuHisValProAsnSerArgSerGlnGlySerValLeuSerCys 323
Qy 61 TGTGGCTCCAGTTCCGGAACAGTAAACCTGTTCCGACTACTGCTCTCCCATATGCTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerGluHisCysLeuPheHisIleVal 343
Qy 121 AATCTTCTCGAGGACTGGGACCTCCGACCGAACATCGAGAACACATCAGGATTCCT 180
Db 344 AsnLeuIleAspAspTrpGlyProCysAlaGluHisGlyGluHisArgThrPro 363
Qy 181 AGGACCTCTGCTGTGTACAGCGGGGTTTTCTCTGTTGACAAGATCTCTCACAATACC 240
Db 364 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 383
Qy 241 GCAGAGCTAGACTCTG-GTGGACTCTCTCAATTTCTAGGGGAGCACCACCTGTC 299
Db 384 SerGluSerArgLeuValAlaAspPheSerGlnPheSerArgGlyAsnThrArgVal-Se 403
Qy 300 CTGGCCAAAATTCGACGTCCCAACCTCCCAATCACTCACCACCTCTTGTCTCCAAATTT 359
Db 403 rTTrpLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAspLe 423
Qy 360 GTCCTGCTATCGCTGATGTGTCTGGGCGTTTTTATCATATTTCTCTTCTCTGCTGTC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl 443
Qy 420 TATCCCTCATCTCTCTGTTGTTCTCTGGAATCAAGGTATGTTCGCTGCTCTCTCTCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 463
Qy 480 TACTTCCAGGAACATCAACACACGACGCGGGCCATCAAGACCTCGACGCTCTGCTGTC 539
Db 463 rAsnSerArgIleIleAsnHisGlnHisArgThrMetGlnAsnLeuHisAspSerCysSe 483
Qy 540 AAGAAACTCTACGTTTCCCTCTTGTGCTGTATCAAAACCTTCGAGCGGAACTGCATTT 599
Db 483 rArgAsnLeuTyrValSerLeuMetLeuLeuTyrLysThrTyrGlyArgLysLeuHisLe 503
Qy 600 GTATTTCCCATCCCATCTCTGCTGCTTTCGAAAGATTCCTATGGAGTGGGCTCAGTCC 659
Db 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Qy 660 GTTTCCTCTGCTCAGTTTACTAGTGCATTTTGTTCAGTGGTTCGTAGGGCTTTCCCCCA 719
```



```

Db      532 sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLeuThrValHisLe 552
QY      780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTGTTCTTGGGTATACATTAACCC 839
Db      552 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyHisLeuAsnPr 572
QY      840 TAATAAACCAAAAGTTGGGGCTACTCCCTTAACCTCATGGATATGTAATTCGAAGTTG 899
Db      572 oAenLysThrLysArgTrpGlyTyrSerLeuHisPheMetGlyTyrValIleGlyCysTy 592
QY      900 GGGTACTTTACCGCAGGAACATATGTTAATAAATCAAGCAATGTTTTCGAAAATGCC 959
Db      592 rGlySerLeuProGlnAspHisIleIleGlnLysLeuGluCysPheArgLysLeuPr 612
QY      960 TGTAAATAGACCTATTGATTTGGAAGTATGTCCTCAAGAAATGTCGGCTTTTGGCTTTC 1019
Db      612 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 632
QY      1020 TGCCCTTTTACAAATGCTGCTATCTGCTTGTATGCTTGTATGCTTGTATGCTATGCTATGCT 1079
Db      632 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 652
QY      1080 TAAGCAGCTTTTCACTTTCTCGCAACTTACAAAGCCCTTTCTGTGTAAACAATATCTGAA 1139
Db      652 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 672
QY      1140 CTTTATCCCTGTCGCGCAACCGTCCGCTCTGCAAGTGTCTGCAAGTGTCTGCAAGCAACCC 1199
Db      672 nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 692
QY      1200 CACTGGATGGGCTTGGCCATGAGCCATCAGCGATGCTGGAACCTTTCTGGCTCTCT 1259
Db      692 oThrGlyTrpGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheGlnAlaProLe 712
QY      1260 GCCCATCCATCTCGGAACCTCTAGCAGCTTTGTTGCTCGCAGCGGCTCGAGCAAA 1319
Db      712 uProIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaAs 732
QY      1320 ACTTATCGGAACCGCAACTCTGTTGCTCTCTCGGAATPACACCTCTCTTCCATGGCT 1379
Db      732 nIleLeuGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 752
QY      1380 GCTAGGTGCTCCCACTGGATCTCGCGGACGCTCTTGTCTACGTCGCTCGCTCGGC 1439
Db      752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 772
QY      1440 GCTGAATCCCGCGACGACCGCTCTCGGGCGGTTTGGGCTCTACCGTCCCTTCTTCA 1499
Db      772 aLeuAsnProAlaAspAspProSerArgLysArgLysLeuSerArgProLeuLeuArg 792
QY      1500 TCTGCGTCTCCGCGCAGCACCGGCGCACCTCTCTTTACGCGTCTCCCGATATGTC 1559
Db      792 gLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValPr 812
QY      1560 TTCTCATCTCGGACCGTGTGCACTTCTGCTTCACTCTGACCTCGCATCGCATGAGACCA 1619
Db      812 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 832
QY      1620 G 1620
Db      832 o 832

```

RESULT 13

S20757
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype ayw, patient E)
C:Species: hepatitis B virus, HBV
A:Variety: subtype ayw, patient E
C:Date: 20-Feb-1995 #sequence_revision 22-Nov-1996 #text_change 09-Jul-2004
C:Accession: S20757
R:Iai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Melis, A.; Porru, A.
submitted to the EMBL Data Library, March 1992
A:Description: Sequence analysis of HBV genomes isolated from patients with HBsAg negati

A:Reference number: S20745
A:Accession: S20757
A:Molecule type: DNA
A:Residues: 1-832 <LAI>
A:Cross-references: UNIPROT:O67882; UNIPARC:UPI00000F25B1; EMBL:X65259; NID:G59439; PIDK:
A:Experimental source: subtype ayw, patient E
C:Genetics:
C:Gene: P
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores: 4.38e-187 Length: 832
Pred. No.: 2637.00 Matches: 494
Score: 95.19% Conservative: 21
Percent Similarity: 91.31% Mismatches: 25
Best Local Similarity: 44.64% Indels: 2
Query Match: 1 Gaps: 0
DB:

US-10-761-006a-1 (1-3215) x S20757 (1-832)

```

QY      1 CTCCCAACATTCACCAAGCTCTGCTAGATCCAGGCTGAGGGGCTATATTTCTCTGC 60
Db      293 LeuHisAsnLeuProAsnSerAlaAArgSerGlnSerGluArgProValPheProCys 312
QY      61 TGGTGGCTCCAGTTCGCGAACAGTAAACCTGTTCGAGCTACTGTCTCTCCCATATGTC 120
Db      313 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerLeuVal 332
QY      121 AATCTTCTCGAGACTGGGGACCTCGACCGAATCATGGAGAACACACATCATGATTCCT 180
Db      333 AsnLeuArgGluAspTrpGlyProCysThrGluHisGlyGluHisIleArgIlePro 352
QY      181 AGACCCCTGCTCGTGTACAGCGGGGTTTCTCGTGACAGAATCTCTCAATACC 240
Db      353 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 372
QY      241 GCAGAGTCTAGACTCTG-GTGAGCTTCTCTCAATTTTCTAGGGGGAGCACCCACGTGTC 299
Db      373 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnTyrArgVal-Se 392
QY      300 CTGGCCAAATTCGAGTCCCGACCTCCCAATCATCATCAACCTCTTGTCTCCCAATTT 359
Db      392 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 412
QY      360 GTCTGGCTATCTCGATGTGTCTCGGGGTTTATCATATCTCTCTCATCTCTGCTGC 419
Db      412 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl 432
QY      420 TATGCTCATCTTCTTGTGTTCTTCTGAGCTACCAAGGTATGTTCCCGTTTGTCTCTC 479
Db      432 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 452
QY      480 TACTTCCAGGAACATCAACCAACGACGCGGGCCATGCAAGACTGCAGCTCTCTGCTC 539
Db      452 rAsnSerArgIlePheAsnAsnGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 472
QY      540 AAGAAACTCTAGCTTTCCTCTTGTGTCTGTACAAAACCTTCGACGGAACCTGCACTT 599
Db      472 rArgAsnLeuTyrValSerLeuLeuLeuTyrGlnThrPheGlyArgLysLeuHisLe 492
QY      600 GTATTCCCATCCCATCATCTCTGCGCTTTCGCAAGATTCTTATGGAGTGGGCTCAGTCC 659
Db      492 uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 512
QY      660 GTTTCCTCTGGCTCAGTTTACTAGTGCATTTGTTTCAGTGGTTCGTAGGGCTTTCCGCC 719
Db      512 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHis 532
QY      720 CTGTTTGGCTTTTCAAGTTATATGATGATGTGTTTGGGGCGGAGTCTGTACACATCT 779
Db      532 sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe 552

```



```

QY 900 GGCTACTTTACCGAGGAACATATTGTACTAAACTCAAGCAATGTTTTCGAAAACTGCC 959
Db ||||| ||||| ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
603 pGlyThrTrpProGlnAspHisIleValGlnAenPheLeuCysPheArgLysLeuPr 623
QY 960 TGTAAATAGACCTATTGATTGGAAGATGTCAAAGAATTGTGGGTCTTTTGGGCTTTGC 1019
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGCCCTTTTACACAATGTGGCTATCCTGCTTGATGCGCTTTATATGCATGTATACAATC 1079
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ::|||
643 aAlaProPheThrGlnCysGlyTyProAlaLeuMetProLeuTyAlaCysIleGlnAl 663
QY 1080 TAAGCAGCCTTTCACCTTCTCGCAACTTACAAAGCCTTCTCTGTAAACAATATCTGAA 1139
Db :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ::|||
663 aLysGlnAlaPheThrPheSerProThrTyLysAlaPheLeuSerLysGlnTyMetTh 683
QY 1140 CCTTTACCCCGTTGCCCGCAAGCGTCTCTGCCAAGTCTTTGCTGACGCAACCCCC 1199
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
683 rLeuTyProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 703
QY 1200 CACTGGATGGGCTTGGCCATAGCCATAGCGCATGGCTGGAACTTTCTGGCTCCTCT 1259
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
703 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValSerProLe 723
QY 1260 GCCGATCCATAGTCGGAACCTCTAGCAGCTTCTTTGCTCGCAGCCGGTCTGGAGCAMA 1319
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
723 uProIleHisThrAlaGluLeuAlaIleCysPheAlaArgSerArgSerGlyAlaAs 743
QY 1320 ACTTATCGGAACCGACAACCTCTGTCTCTCTCGGAATACACCTCTCTTCCATGGCT 1379
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
743 nLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyThrSerPheProTrpLe 763
QY 1380 GCTAGGTGTGCTGCCAATCGGATCCTCGCGGAGCGTCTTTGTCTACGTCCTCGTCGGC 1439
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyValProSerAl 783
QY 1440 GCTGAATCCCGGAGCAGCCGCTCGGGCGCTTTGGGGCTCTACGTCCTCGCTCTTCA 1499
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyArgProLeuLeuAr 803
QY 1500 TCTGCGGTTCCGGCCGACACCGGGCGCACCTCTCTTTACGGCTCTCCCGTATGTGCC 1559
Db ||||| ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
803 gLeuProTyArgProThrThrGlyArgThrSerLeuTyAlaAspSerProSerValPr 823
QY 1560 TTCTCATCTCCCGGACCGGTGTGACCTTGCCTTCACTCTGACGTGCGATGGAGACACC 1619
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
QY 1620 G 1620
Db |
843 o 843

```

Search completed: December 2, 2005, 00:02:30
Job time : 160 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 1, 2005, 23:14:50 ; Search time 706.5 Seconds

(without alignments)
6421.156 Million cell updates/sec

Title: US-10-761-006A-1

Perfect score: 5907

Sequence: 1 CTCACACACATTCACACAG.....CCTCAGGCCACGAGTGGAA 3215

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool.p/US10761006/runat.01122005.113952.26869/app.query.fasta.1.3399
-DB=UniProt -OPMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORESPCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10761006.CGN 1.1199 @runat.01122005.113952.26869 -NCPU=6 -ICPU=3
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 05.80:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4130	69.9	843	Q69616_HPBVO	Q69616 hepatitis b
2	3745.5	63.4	832	Q67907_HPBVO	Q67907 hepatitis b
3	3693.5	62.5	842	Q69590_HPBVO	Q69590 hepatitis b
4	2824	47.8	843	Q7D85_HPBVO	Q7D85 hepatitis b
5	2822	47.8	843	Q4FD83_HPBVO	Q4FD83 hepatitis b
6	2810	47.6	540	Q69026_HPBVO	Q69026 hepatitis b
7	2808	47.5	843	Q69045_HPBVO	Q69045 hepatitis b
8	2807	47.5	843	Q69045_HPBVO	Q69045 hepatitis b
9	2806	47.5	540	Q69045_HPBVO	Q69045 hepatitis b
10	2806	47.5	843	Q69045_HPBVO	Q69045 hepatitis b
11	2804	47.5	843	Q69045_HPBVO	Q69045 hepatitis b
12	2803	47.5	843	Q69045_HPBVO	Q69045 hepatitis b
13	2803	47.5	843	Q69045_HPBVO	Q69045 hepatitis b
14	2802	47.4	838	Q69045_HPBVO	Q69045 hepatitis b
15	2802	47.4	843	Q69045_HPBVO	Q69045 hepatitis b
16	2802	47.4	843	Q69045_HPBVO	Q69045 hepatitis b

17	2801	47.4	843	2	Q80MR4_HPBVO	Q80MR4 hepatitis b
18	2800	47.4	843	2	Q5DVZ8_HPBVO	Q5DVZ8 hepatitis b
19	2799	47.4	843	2	Q5K35_HPBVO	Q5K35 hepatitis b
20	2799	47.4	843	2	Q4FD61_HPBVO	Q4FD61 hepatitis b
21	2798	47.4	843	2	Q91549_HPBVO	Q91549 hepatitis b
22	2798	47.4	843	2	Q80H07_HPBVO	Q80H07 hepatitis b
23	2798	47.4	843	2	Q5KR19_HPBVO	Q5KR19 hepatitis b
24	2797	47.4	843	2	Q6YLMO_HPBVO	Q6YLMO hepatitis b
25	2796	47.3	843	2	Q80MQ8_HPBVO	Q80MQ8 hepatitis b
26	2796	47.3	843	2	Q8JVC9_HPBVO	Q8JVC9 hepatitis b
27	2795	47.3	843	2	Q5SDK1_HPBVO	Q5SDK1 hepatitis b
28	2794	47.3	843	2	Q9QNS4_HPBVO	Q9QNS4 hepatitis b
29	2794	47.3	843	2	Q9YZR5_HPBVO	Q9YZR5 hepatitis b
30	2794	47.3	843	2	Q8BCI7_HPBVO	Q8BCI7 hepatitis b
31	2792	47.3	843	2	Q91547_HPBVO	Q91547 hepatitis b
32	2792	47.3	843	2	Q99217_HPBVO	Q99217 hepatitis b
33	2792	47.3	843	2	Q9QM17_HPBVO	Q9QM17 hepatitis b
34	2792	47.3	843	2	Q9QMJ8_HPBVO	Q9QMJ8 hepatitis b
35	2791.5	47.3	844	2	Q4FD77_HPBVO	Q4FD77 hepatitis b
36	2791	47.2	540	2	Q97976_HPBVO	Q97976 hepatitis b
37	2791	47.2	843	2	Q39671_HPBVO	Q39671 hepatitis b
38	2791	47.2	843	2	Q81134_HPBVO	Q81134 hepatitis b
39	2791	47.2	843	2	Q8VIM4_HPBVO	Q8VIM4 hepatitis b
40	2790	47.2	843	2	Q91541_HPBVO	Q91541 hepatitis b
41	2790	47.2	843	2	Q5SDK5_HPBVO	Q5SDK5 hepatitis b
42	2790	47.2	843	2	Q5KR43_HPBVO	Q5KR43 hepatitis b
43	2790	47.2	843	2	Q5R219_HPBVO	Q5R219 hepatitis b
44	2790	47.2	843	2	Q9QMK4_HPBVO	Q9QMK4 hepatitis b
45	2789	47.2	843	2	Q9PWT8_HPBVO	Q9PWT8 hepatitis b

ALIGNMENTS

RESULT 1

Q69616_HPBVO
ID Q69616_HPBVO PRELIMINARY; PRT; 843 AA.
AC Q69616;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polymerease (Fragment).
GN Name=P;
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OK NCBI_taxid=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Genotype C;
RX MEDLINE=94120723; PubMed=8291231;
RA Norder H., Courouce A.M., Magnius L.O.;
RT "Complete genomes, phylogenetic relatedness, and structural proteins
of six strains of the hepatitis B virus, four of which represent two
new genotypes.";
RL Virology 198:489-503 (1994).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Genotype c;
RA Norder H.M.L.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; X75656; CAA53338.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004523; F:Ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001462; DNAPol_viral_C.
DR InterPro; IPR000477; DNAPol_viral_N.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.

DR Pfam: PF00078; RVT_1; 1;
DR ProDom: PD00814; DNapi_viral_C; 1.
DR PROSITE: PS00878; RT_POL; 1.
FT CHAIN 541 588 core antigen.
FT NON_TER 1 1
FT NON_TER 843 843
SQ SEQUENCE 843 AA; 94603 MW; 934BEEC1F4E235D0 CRC64;

Alignment Scores:
Pred. No.: 1 55e-299 Length: 843
Score: 4130.00 Matches: 802
Percent Similarity: 76.26% Conservative: 17
Best Local Similarity: 74.67% Mismatches: 24
Query Match: 69.92% Indels: 232
DB: 2 Gaps: 1

US-10-761-006A-1 (1-3215) x Q69616_HPBVO (1-843)

QY 1 CTCCACAACATCCACCAAGCTCTGCTAGATCCAGGGTGAGGGGCTATATTTCTCTGC 60
DB 1 LeuHisAsnIleProThrSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 20
QY 61 TGTGGCTCCAGTTCGGGAACAGTAACCTGTTCGACTACTGCTCTCCCATATCGTC 120
DB 21 TrpTrpLeuGlnPheArgAnSerLysProCysSerAspTyrCysLeuSerHisPheVal 40
QY 121 AATCTTCTCAGACATGGGGACCTCCACCGAACATGGAGAACACATCAGGATTCCT 180
DB 41 AsnLeuLeuGluAspTrpGlyProCysAsnGluHisGlyGluHisAsnIleArgIlePro 60
QY 181 AGGACCTCTCGTGTATACAGCGGGGTTTTCTCGTTGACAGAAATCTCAATACC 240
DB 61 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 80
QY 241 GCAGAGCTAGACTCG-GTGGACTTCTCTCAATTTCTAGGGGAGCACCCACGGTTC 299
DB 81 ThrGluSerArgLeuGlyValAspPheSerGlnPheSerArgGlySerThrLysVal-Se 100
QY 300 CTGGCCAAATTCGAGTCCCAACCTCCAATCACTCACCAACCTCTGTCTCTCCAAATTT 359
DB 100 xTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 120
QY 360 GTCTCGCTATCGCTGGATGTCTCGGGGTTTTATCATATTCCTCTTCATCTCTCTGC 419
DB 120 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl 140
QY 420 TATGCTCATCTTCTGTGTCTCTGACTTACCAAGTATGTTGCCGTTTGTCTC 479
DB 140 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 160
QY 480 TACTTCCAGGAACATCAACACAGACGCGGGCCATGCAAGACCTGCACGACTCTCTGCTC 539
DB 160 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 180
QY 540 AAGAAACTCTAGCTTCTCTGTGTGTGTACAAACCTTCGAGCGGAACCTGACTT 599
DB 180 rArgAsnLeuTyrValSerLeuMetLeuTyrLysThrPheGlyArgLysLeuHisLe 200
QY 600 GTATTCCCATCCATCATCTCTGGGCTTCGCAAGATTCTATGGGAGTGGGCTCAGTCC 659
DB 200 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 220
QY 660 GTTTCCTCTCGCTCAGTTTATAGTCCCATTTGTTTCAGTGGTTCTGAGGGCTTTCCCCCA 719
DB 220 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 240
QY 720 CTGTTTGGCTTTCAGTATATGATCATCTGTTATTTGGGGGGAAGTCTGTACAACATCT 779
DB 240 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 260
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
DB 260 uGluSerLeuTyrThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 280

QY 840 TAATAAAACCAACCGTTGGGCTACTCCTTAACCTTCATGGGATATGTAATTAAGAGTTG 899
DB 280 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 300
QY 900 GGTACTTTTACCGCAGGAACATATTGTACTAAACCTCAAGCAATGTTTTCGAAAACCTGCC 959
DB 300 pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr 320
QY 960 TGTAAATAGACCTATTGATTGAAAGTATGTCAAAGAATTGTGGGTCTTTTGGGCTTTGC 1019
DB 320 oValAsnArgProIleAsnTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 340
QY 1020 TGCCCTTTTACCAATGTGGCTATCTCCTCTGTATGCTTATATGATGATATACAATC 1079
DB 340 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnAl 360
QY 1080 TAAGCAGGCTTTCACCTTCTGCCCACTTACAGGCTTCTGTGTAACAATATCTGAA 1139
DB 360 aLysGlnAlaPheThrPheSerSerThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 380
QY 1140 CCTTTACCCCGTTGCCGCAACGGTCCGTCTCTGCCAAGTGTGTTGCTGACGCAACCCC 1199
DB 380 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 400
QY 1200 CACTGATGGGCTTGGCCATAGGCCATCAGCGCATGGCTGGAACCTTTCTGGCTCTCT 1259
DB 400 oThrGlyTrpGlyLeuAlaIleGlyAsnGlnArgMetArgGlyThrPheValAlaProLe 420
QY 1260 GCGATCCATATCTGCGAATCTCTAGCAGCTTGTGTTGCTCGCAGCGGCTCTGGAGCAA 1319
DB 420 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy 440
QY 1320 ACTTATCGGAACCGCAACTCTGTGTCTCTCTCGGAATATACACTCTCTTCCATGGCT 1379
DB 440 sleuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 460
QY 1380 GTAGGCTGTCTGCCAATCTGCGCGGAGCTCTTGTGTCTACGTCCGCTCGCG 1439
DB 460 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 480
QY 1440 GTGAATCCCGGGAGCAGCCCTCTCGGGCGGTTTGGGGCTCTACCGTCCCTTTCTTCA 1499
DB 480 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyIleTyrArgProLeuLeuAr 500
QY 1500 TCTGCGCTTCCGCGCGCACCCACCGCGGCTCTCTTTACGGGCTCTCCCGTATGTGCC 1559
DB 500 gLeuArgPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 520
QY 1560 TTCTCATCTGCGGACCGGTGTCACCTTCGCTTACCTCTGCACTGCGCATGGAGACACC 1619
DB 520 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 540
QY 1620 GTGAACGCAACCGCAGGTCTTGCCCAAGGTCTTATATAAGAGGACTCTTGAGACTCTCAGCA 1679
DB 540 o----- 540
QY 1680 ATGTCAACGACGACCTTGAGGCATCTTCAAAGACTGTGTGTTTAAAGACTGGGAGGAG 1739
DB 540 ----- 540
QY 1740 TTGGGGGAGGAGATTAGGTTAAAGATTATATGTAAGAGGCTGTAGGCAATAAATGGTTC 1799
DB 540 ----- 540
QY 1800 TGTTTACCAGCACCACGAACTTTTCTCCTCTGCGCTAATCATCTCATGTTTCATGTCCTA 1859
DB 540 ----- 540
QY 1860 CTGTTTCAAGCTTCCAAAGTGTGCTTGGGTGGCTTTGGGACATGGACATTGACCCGTATA 1919
DB 540 ----- 540

QY 1920 AAGNAITTTGGAGCATCTGCTGAGTACTCTCTTTTGGCCCTCTGACTTCTTCTCCGCTCA 1979
 Db 540 ----- 540
 QY 1980 TTGAGATCTCTCGACACCGCTCTGCTCTGTATCGGAGGCGCTTAGAGTCTCCGGAAC 2039
 Db 540 ----- 540
 QY 2040 ATTGTTCCCTCACCATACAGCACTCAGGCAAGCTATTTGTGTGGGTGAGTTGATGA 2099
 Db 540 ----- 540
 QY 2100 ATCTGGCCACCTGGGTGGGAAGTAATTGGAAGATCCAGCATCCAGGAATTAGTAGTCA 2159
 Db 540 ----- 540
 QY 2160 GCTATGTCAACGTTAATATATGGGCTAAACTCAGACAAATATTGTGTTTCACATTTCT 2219
 Db 540 ----- 540
 QY 2220 GTCTTACTTTTGAAGAGAACTGTTCTTGAGTACTTGGTATCTTTTGGAGTGTGGATTC 2279
 Db 540 ----- 540
 QY 2280 GCACCTCTACGCTTACAGACCACCAATGCCCTATCTTATCAACACTTCCGGAACACTA 2339
 Db 541 -----Met--ProLeuSerTyrGlnHisPheArgLysLeu 551
 QY 2340 CTGTTGTTAGACGAGGAGGCTCCCTAGAGAAGAACTCCCTCGCTCGGACAGCAA 2399
 Db 552 LeuLeuLeuAspValGluAlaGlyProLeuGluGluGluLeuProArgLeuAlaAspGlu 571
 QY 2400 GGTCCTCAATCGCCCGTGGCGAAGATCTCAATCTCGGATCTCAACGTTAGTATTCCT 2459
 Db 572 GlyLeuAsnHisArgValAlaGluAspLeuAsnLeuGlyAsnProAsnValSerIlePro 591
 QY 2460 TGGACTATAAGTGGGAACCTTACTGGGCTTATTCTTCTACTGACTGCTCTTAAAT 2519
 Db 592 TrpThrHisLysValGlyAsnPheThrGlyLeuTyrSerSerThrValProValPheAsn 611
 QY 2520 CCCGAGTGGCAAAATCTCTCTTCCCTCACATTCATTACAGAGGACATTTATTAATPAGA 2579
 Db 612 ProGluTrpGlnThrProSerPheProAspIleHisLeuGlnGluAspIleIleAsnArg 631
 QY 2580 TGTCACAAATATGCGGCCCTCTTACAGTTAATGAAAAGAAAGATTAATAATTAATG 2639
 Db 632 CysGlnGlnTyrValGlyProLeuThrValAsnGluLysArgLeuLysLeuIleMet 651
 QY 2640 CCTGCTAGTTTATCTTAACTTAAATATTTGCTTACAGCAAGGACATTAACCG 2699
 Db 652 ProAlaArgPheTyrProAsnLeuThrLysTyrLeuProLeuAspLysGlyIleLysPro 671
 QY 2700 TATTATCTGAACATGCAAGTTAATCACTTACTTCAAACTAGGCATTATTTACATACTCTG 2759
 Db 672 TyrTyrProGluHisAlaValAsnHisTyrPheLysThrArgHisTyrLeuHisThrLeu 691
 QY 2760 TGGAGGCTGCATCTTATATAAGAGAAACTACAGCGCGCTCATTTTGTGGGTCA 2819
 Db 692 TrpLysAlaGlyIleLeuTyrLysArgGluThrThrArgSerAlaSerPheCysGlySer 711
 QY 2820 CCATATCTTGGACACAGACTACAGATCGGAGTTGTGCTTCCAACTCCGACAGG 2879
 Db 712 ProTyrSerTrpGlnGlnLeuGlnHisGlyArgLeuValPheGlnThrSerGluArg 731
 QY 2880 CATGGGAGCACTTGTGTTTCCCAATCTCTCGGATCTTTTCCCGATCCACGATTTGGA 2939
 Db 732 HisGlyAspGluSerPheCysSerGlnSerSerGlyIleLeuSerArgSerProValGly 751
 QY 2940 CCTGCGTTCCGAGCAACTCAAACTCAAGTTGGAGTTCACCCCAACAGGATCA 2999
 Db 752 ProCysIleArgSerGlnLeuLysGlnSerArgLeuGlyLeuGlnProGlnGlnGlySer 771
 QY 3000 CTGGCAGGCAATCAGTAGGATGGGAGCATTCGGGCCAGGTTTACCCACCCACACA 3059

Db 772 MetAlaArgGlyLysSerGlyArgSerGlySerIleArgAlaArgValHisProThrThr 791
 QY 3060 CGCGGCTCTTTGGGGGAGCCCTCAGCTCAGGCGATATTTCACACAGTCCGAGCAGC 3119
 Db 792 ArgArgSerPheGlyValGluProSerGlyThrArgHisIleAspAsnHisAlaSerSer 811
 QY 3120 ACCTCTCTCTCGCTCCACCAATCGGAGTCCAGGACAGCAGCTACTCTCCATCTCTCCACC 3179
 Db 812 SerSerSerCysLeuHisGlnSerAlaValArgLysThrAlaTyrAlaHisLeuSerThr 831
 QY 3180 TCTAAGACAGCATCTCTCAGGCCACGCGAGTGGAA 3215
 Db 832 SerLysArgGlnSerSerGlyHisAlaValGlu 843
 RESULT 2
 Q67907_HPBVO PRELIMINARY; PRT; 832 AA.
 ID Q67907;
 AC Q67907;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE DNA polymerase (Fragment).
 OS Hepatitis B virus.
 OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
 OC Orthohepadnavirus.
 OC NCBI_TaxID=10407;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Sub-type ayw;
 RX MEDLINE=94079539; PubMed=8257295;
 RA Preisler-Adams S., Schlayer M.J., Peters T., Hettler F., Gerok W.,
 Rasenack J.;
 RA "Sequence analysis of hepatitis B virus DNA in immunologically
 RT negative infection.";
 RL Arch. Virol. 133:385-396(1993).
 CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
 DR EMBL; X72702; CAA51254.1; -; Genomic_DNA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0004523; F:ribonuclease H activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
 DR InterPro; IPR001462; DNAPol_viral_C.
 DR InterPro; IPR00201; DNAPol_viral_N.
 DR InterPro; IPR00477; RTase.
 DR Pfam; PF00336; DNA_pol_viral_C; 1.
 DR Pfam; PF00242; DNA_pol_viral_N; 1.
 DR Pfam; PF00078; RVT_1; 1.
 DR ProDom; PD000814; DNAPol_viral_C; 1.
 DR PROSITE; PS00878; RT_POL; 1.
 FT NON TER 1
 FT NON TER 832
 SQ SEQUENCE 832 AA; 93864 MW; EC6872097F04BAAF CRC64;
 Alignment Scores:
 Pred. No.: 8,42e-271 Length: 832
 Score: 3745.50 Matches: 726
 Percent Similarity: 72.16% Conservative: 49
 Best Local Similarity: 67.60% Mismatches: 57
 Query Match: 63.41% Indels: 243
 DB: 2 Gaps: 2
 US-10-761-006a-1 (1-3215) x Q67907_HPBVO (1-832)
 QY 1 CTCACCAACATTTCCACCAAGCTCTGTAGATCCAGGCTGAGGGCCCTATATTTCTCTGC 60
 Db 1 LeuHisAsnPheProProAsnSerAlaArgSerGlnGlyGluArgProValPheProCys 20
 QY 61 TGTGGCTCCAGTTCCGGAAACAGTAACCCCTGTTCGCACTACTGCTCTCCCATATCGTC 120
 Db 21 TrpTrpLeuGlnPheArgAsnSerLysProCysSerLysProCysLeuSerHisIleVal 40

121	QY	121	AATCTTCTCGAGACTGGGACCTGCACCGAA	CATGGAGAA	CAACAACATCAGGATTCCT	180
41	Db	41	AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisIleAspGlePro			60
181	QY	181	AGGACCCCTGCTCGTGTATACAGGGGGTCTTCTCGTTCACAGAATCCTCAATACC			240
61	Db	61	ArgThrProAlaArgValThrGlyValPheLeuValAspIysAsnProHisAsnThr			80
241	QY	241	GCAGAGTCTAGACTCTG- GTGGAGCTTCTCTCAATTTTCTAGGGGGAGCACCACGCTGTTCT			299
81	Db	81	AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnHisArgVal- Se			100
300	QY	300	CTGGCCAAAATTTCGAGTCCCAACCTCCAATCACTACCAACCTCTTGTCTCCCAATT			359
100	Db	100	rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe			120
360	QY	360	GTCTCGGCTATCGCTGGATGTGTCTGGCGGTTTTATCATATCTCTCTCATCTGCTGC			419
120	Db	120	uSerTrpLeuSerLeuAspValSerAlaIaPheTyHisLeuProLeuHisProAlaSe			140
420	QY	420	TATGCCTCATCTTCTTGTGTTCTTGACACTACCAAGTATGTGCCCGTTGTGCTCT			479
140	Db	140	rMetProHisLeuLeuValGlySerThrGlyLeuSerArgTy-ValAlaArgValSerSe			160
480	QY	480	TACTTCCAGGAAATCAACACACGACGCGGGCCATCGACAGACTCGACACACTCTGCTC			539
160	Db	160	rAsnSerArgIlePheAsnHisGlnArgGlyThrMetGlnAsnLeuHisAspTyCysSe			180
540	QY	540	AAGGAAACTCTAGCTTTCCCTCTTGTGTGTACAAAACCTTCGGACGGAAACTGCAC			599
180	Db	180	rArgAsnLeuTy-ValSerLeuLeuLeuLeuTyGlnThrPheGlyArgIysLeuHisLe			200
600	QY	600	GTATTCATCCATCATCTCGGCTTCCGACAGATTCCTATGGAGTGGGCTCAGTCC			659
200	Db	200	uTySerHisProIleIleLeuGlyPheArgIysIleProMetGlyValGlyLeuSerPr			220
660	QY	660	GTTTCTCTCGCTCAGTTTACTAGTGCAATTGTGTAGTGCTTCGTAGGCTTTCC			719
220	Db	220	oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProH			240
720	QY	720	CTGTTTGGCTTTT CAGTTATATGATGATGTGTTATGGGGCGCAAGTCTGTACAACATCT			779

Qy	1200	CAC	TGATGGGGCTTGGCCATAGGCCCATAGCCCATAGCGCATTGGTGGAAACCTTTCTGGCTCTCT	1259
Db	400	oThrGlyTrpGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheLeuAlaProLe	:::::	420
Qy	1260	GCGCACTCCATACTGCAGAACCTCTACAGACTTCCTTTTTGTCTGCGCAGCGGTCTCGAGCAA	A	1319
Db	420	uProlIleHisThrAlaGluLeuLeuAlaalaCysPheAlaAargSerArgSerGlyAlaAs		440
Qy	1320	ACTTATCGGAACCGACAACCTCTGTTGTCCTCTCTCGGAATAACACCTCTCTTCCATGGCT		1379
Db	440	nIleLeuGlyThrAspAsnSerValValLeuSerArgLysTyThrSerPheProTrpLe		460
Qy	1380	GCTAGGGTGCTGCCAACCTGGATCTCTGCGCGGAGCTCTCTTGTCTACGTCCTCCGTCGGC		1439
Db	460	uLeuGlyCysAlaAlaAsnTrpIleLeuAargGlyThrSerPheValTyValProSerAl		480
Qy	1440	GCTGAATCCCGCGAGCACCGCTCTCGGGGCCGTTTTGGGCTCTACCCTCCCCCTTCTCA		1499
Db	480	aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyIlePheArgProLeuLeuAr		500
Qy	1500	TCTGCCGTTCCGGCCGACACGGGGCGCACCTCTCTTTAGCGGCTCTCCCGGTATGTGCC		1559
Db	500	gLeuProPheArgProThrThrGlyArgThrSerLeuTyAlaAspSerProSerValPr		520
Qy	1560	TTCTCATCTCCCGGACCGTGTGCACCTTCGTTTCCACCTCTGCACGTGCGCATGGAGACCAC		1619
Db	520	oSerHisLeuProValArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr		540
Qy	1620	GTGAACGACGCGCAGGTCTTGGCCCAAGTCTTATATAAGAGGACTCTTGGACTCTCAGCA		1679
Db	540	o-----		540
Qy	1680	ATGTC AACGACCGACCTTGAGGCATCTTCAAAGACTGTGTGTTTAAAGACTGGAGGAG		1739
Db	540	-----		540
Qy	1740	TTGGGGAGGAGATTAGGTTTAAAGATTTATGTACTAGGAGGCTGTAGGCATAAATTGGTC		1799
Db	540	-----		540
Qy	1800	TGTTACACGACGATGCAACTTTTTTCTCTCTGCGCTTAATCATCTCATGTTCTATGCTCTA		1859
Db	540	-----		540
Qy	1860	CTGTTCAAGCCTCCAAGCTGTGCTTGGGTGGCTTTTGGGACATGGACATTTGACCCGTATA		1919
Db	540	-----		540
Qy	1920	AAGAAATTGAGCATCTGCTGAGTTACTCTCTTTTTTTTGGCTTCTGACTTCTTCCGTCCTA		1979
Db	540	-----		540
Qy	1980	TTGAGATCTCTCGACACCGCTCTGCTCTGTATCGGAGGCGTTAGAGTCTCCGGAAC		2039
Db	540	-----		540
Qy	2040	ATTGTTGCGCTCAACATACAGCACTCAGGCAAGCTATTTTGTGTGGGGTGAGTTGANGA		2099
Db	540	-----		540
Qy	2100	ATCTGGCCACTGGGTGGGAAGTAATTGGGAAGNTCCAGCATCCAGGGAATTAGTAGTCA		2159
Db	540	-----		540
Qy	2160	GCTATGTCAACGTTAATATATGGGCTAAAACTCAGACAAATATTGTGTTTTCACATTTCCCT		2219
Db	540	-----		540
Qy	2220	GTCTTACTTTTGAAGAGAAAACCTGTTCTTGAGTACTTGGTATCTTTTGGAGTGTGGATTCT		2279
Db	540	-----		540
Qy	2280	GCATCTCTACGCTTACAGACACCCAAATGCCCTATCTTTATCAACACTTTCCGGAACACTA		2339

```
Db 541 -----Met--ProLeuSerTyrGlnHisPheArgLysLeu 551
Qy 2340 CTGTTGTTAGACGACGAGCGAGTCCCTAGAGAGAACTCCCTCGCTCGCAGACGAA 2399
Db 552 LeuLeuLeuAapGluAaGlyProLeuGluGluLeuProArgLeuAaAaspGlu 571
Qy 2400 GGTCTCAATCGCCGCTCGCAGAGATCTCAATCTCGGGAATCTCAACGTTAGTATTCCT 2459
Db 572 AspLeuAenArgArgValAaGluAaspLeuAenGlyAsnLeuAenValSerIlePro 591
Qy 2460 TGGACTCATAGGTGGGAACTTACTGGGCTTTATTCTTCTACTGTACCTGCTTTAAT 2519
Db 592 TrpThrHisLysValGlyAsnPheThrGlyLeuTyrSerSerThrValProValPheAsn 611
Qy 2520 CCCGAGTGGCAAAATCTCTCTTCCTCCTCATTATTCATTAAGAGACATTAATATAGA 2579
Db 612 ProHisTrpLysThrProSerPheProAsnIleHisLeuHisGlnAaspIleLeLysLys 631
Qy 2580 TGTCAACATATGTGGCCCTCTTACAGTTAATGAAAAGAGAGATTAATAATTATG 2639
Db 632 CysGluGlnPheValGlyProLeuThrValAsnGluLysArgArgLeuGlnLeuIleMet 651
Qy 2640 CCTCTAGGTTTATCTTAACCTTACTAAATATTGCTTCTTACGAAAGCATTAACCG 2699
Db 652 ProAlaArgPheTyrProAsnValThrLysTyrLeuProLeuAapLysGlyLysPro 671
Qy 2700 TATTATCTGAACATGCAATTAATCATTACTTCAAACTAGGCATTAATTACATCTCTG 2759
Db 672 TyrTyrProGluHisLeuValAsnHisTyrPheGlnThrArgHisTyrLeuHisThrLeu 691
Qy 2760 TGGAGGCTGTCATCTATATAAGAGAACTACAGCGCGCTCATTTTGGGTCA 2819
Db 692 TrpLysAaGlyLeuTyrLysArgGluThrThrArgSerAlaSerPheCysGlySer 711
Qy 2820 CCATATTCTTGGAAACAGACTACAGCATGGGAGTGGTCTTCCAAACCTCGACAAAG 2879
Db 712 ProTyrSerTrpGluGlnLeuGln-----720
Qy 2880 CATGGGAGCAATCTGCTGTTCCCAATCTCTGGATTTCTCCGATACCAAGTTGGA 2939
Db 721 HisGlyAlaGluSerPheHisGlnGlnSerGlyIleLeuSerArgProProValGly 740
Qy 2940 CCCTGGCTCGGAGCAACTCAAACTCAGATGGGACTTCAACCCACACAGGATCA 2999
Db 741 SerSerLeuGlnSerLysHisArgLysSerArgLeuGlyLeuGlnSerGlnGlnGlyHis 760
Qy 3000 CTGGCCAGAGCAATCAGGTAGGAGTGGGAGCATTCGGGCGAGGTTTCCACCCACCA 3059
Db 761 LeuAlaArgArgGlnGlnGlyArgSerTrpSerIleArgAlaGlyIleHisProThrAla 780
Qy 3060 CGGCGGTCTTTTGGGGGGAGCCCTCAGGCTCAGGCAATATTGACACAGTGCCAGCAGC 3119
Db 781 ArgArgProPheGlyValGluProSerGlySerGlyHisAsnThrAsnLeuAlaSerLys 800
Qy 3120 ACTCTCTCTGCTCCCAATCGCAGTTCAGGAGACAGGCTTCTCCCTCTCCAC 3179
Db 801 SerAlaSerCysIleTyrGlnSerProValArgLysAlaAlaTyrProAlaValSerThr 820
Qy 3180 TCTAAGACACATCATCTCAGGCCACCGAGTGGAA 3215
Db 821 PheGluLysHisSerSerSerGlyHisAlaValGlu 832
RESULT 3
Q69590 HPBV0
ID Q69590 HPBV0 PRELIMINARY; PRT; 842 AA.
AC Q69590;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polymerase (Fragment).
GN Name=P;
OS Hepatitis B virus.
```

```
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Genotype e;
RX MEDLINE=94120723; PubMed=8291231;
RA Norder H., Courouce A.M., Magnus L.O.;
RT "Complete genomes, phylogenetic relatedness, and structural proteins
of six strains of the hepatitis B virus, four of which represent two
new genotypes."
RL Virology 198:489-503 (1994).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Genotype e;
RA Norder H.M.L.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
EMBL; X75657; CAA53339.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003664; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001462; DNAPol_viral_C.
DR InterPro; IPR000201; DNAPol_viral_N.
DR InterPro; IPR000477; RTase.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00078; RVT_1; 1.
DR ProDom; PD000814; DNAPol_viral_C; 1.
DR PROSITE; PS00878; RT_BOL; 1. Potential.
FT CHAIN <1> >842
FT NON TER 1
FT NON TER 842
FT SEQUENCE 842 AA; 94812 MW; 45C3CD71CF5A4PE7 CRC64;
SQ
Alignment Scores:
Pred. No.: 6,51e-267 Length: 842
Score: 3693.50 Matches: 719
Percent Similarity: 71.85% Conservative: 52
Best Local Similarity: 67.01% Mismatches: 70
Query Match: 62.53% Indels: 233
DB: 2
US-10-761-006A-1 (1-3215) x Q69590_HPBV0 (1-842)
Qy 4 CACAACATTCACCAAGCTCTGCTAGATCCAGGGTGAGGGCCCTATATTTCTGCTGG 63
Db 2 HisAsnIleProProSerSerAlaGlySerGlnSerLysArgProValPheSerCysTrp 21
Qy 64 TGGCTCCAGTTCCGGAACAGTAACCCCTGTTCCGACTATCGCTCTCCCATCGTCAAT 123
Db 22 TrpLeuGlnPheArgAsnSerGluProCysSerAspTyrCysLeuThrHisLeuValAsn 41
Qy 124 CTTCTCGAGGACTGGGACCCCTGCACCGACATCGAGAACACATCAGGATTCCTAGG 183
Db 42 LeuLeuGluAepTrpGlyProCysThrGluHisGlyLysHisIleArgIleProArg 61
Qy 184 ACCCTCTGCTGTTTACAGCGGGGTTTTCTCGTTGACAGAATCTCCACAAATCCCA 243
Db 62 ThrProAlaArgValThrGlyValPheLeuValAapLysAsnProHisAsnThrAla 81
Qy 244 GAGTCTAGACTCTG-GTGGACTTCTCAATTTTCTAGGGGAGACCCACGTTCTCTG 302
Db 82 GluSerArgLeuValValAaspPheSerGlnPheSerArgGlySerArgVal-SerTr 101
Qy 303 GCCAAATTCGACGTCCCACTCCCACTCACTCACCACCTCTTGCTCTCCAAATTTGTC 362
Db 101 pProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSe 121
Qy 363 CTGGCTATCGTGGATGTCTCGGGCGTTTTTATCATATTTCTCTTCATCTCTGCTGCTAT 422
```


|||||
121 rTPrLeuSerLeuAspValSerAlaalaPheTyrHisLeuProLeuHisProAlaIaIaMe 141
|||||
423 GCCTCATCTTCTGGTCTCTCTGAGTACCAAGGTATGTTGCCGTTGTCTCTAC 482
|||||
141 tProHisLeuLeuValGlySerSerGlyLeuSerArglyrValAlaArgLeuSerSerAs 161
|||||
483 TTCCAGGAACATCAACACACGACGGGCCATGCAAGACCTCAGACTCTCTGCTCAAG 542
|||||
161 nSerArgIleAAsnHisGlnTyrGlyThrLeuProAsnLeuHisAspSerCysSerAr 181
|||||
543 GAAACTCTAGCTTCCCTCTGTGTGTGTACAAAACCTTCGGACGGAAACCTGCACTTGTA 602
|||||
181 gAnLeuTyrValSerLeuMetLeuLeuPheLysThrPheGlyArgLysLeuHisLeu 201
|||||
603 TTCCCATCCCATCATCTCGGCTTTCCCAAGATTCCTATGGAGTGGGCTCAGTCCGTT 662
|||||
201 rSerHisProIleWetGlyPheArgLysIleProMetGlyValGlyLeuSerProPh 221
|||||
663 TCTCCTGGCTCAGTTTACTAGTCCCATTTGTTTCAGTGGTTCGTAGGCTTCCCCCACTG 722
|||||
221 eLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHisCy 241
|||||
723 TTTCGGCTTTCAGTTATATGATGATGTGTATTGGGGCGAAGCTCTGTCAACATCTTGA 782
|||||
241 sLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLeuGl 261
|||||
783 GTCCCTTTTACCTTATACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCTAA 842
|||||
261 uSerLeuTyrThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAs 281
|||||
843 TAAACCAACGTTGGGCTACTCCCTTAACCTTCATGGGATGATTAATTGGAGTTGGGG 902
|||||
281 nLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrpGl 301
|||||
903 TACTTTACCGCAGCAATATGCTTAAACCAAGCAATGTTTTCGAAACCTGCTGT 962
|||||
301 ySerLeuProGlnGluHisIleGlnLysIleLysAspCysPheArgLysLeuProVa 321
|||||
963 AAATAGACTATTGATGAAAGTATGTCAAGAATTGTGGGTCTTTTGGGCTTCTCTGC 1022
|||||
321 lAenArgProIleAseTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAl 341
|||||
1023 CCCTTTTACCAATGTGGCTATCTCTGCTGTATGCTTATATGATGATATACAACTAA 1082
|||||
341 aProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysThrGlnSerLy 361
|||||
1083 GCAGGCTTTCACCTTCTCGCCAACTTACAAGGCTTTCGTGTAAACAATATCTGAACCT 1142
|||||
361 sGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsnLe 381
|||||
1143 TTACCCGTTGCCGCAACGGTCCGCTCTCTGCCAAGTGTGTTGTCGACCAACCCCAAC 1202
|||||
381 uTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrProTh 401
|||||
1203 TGGATGGGCTTGGCCATAGGCATCAGCGCATGGCTGGAACCTTCTGGCTCTCTGCC 1262
|||||
401 rGlyTrpGlyLeuAlaIleGlyLeuGlnArgMetArgGlyThrPheValAlaProLeuPr 421
|||||
1263 GATCCATCTCGCGAACTCTACGAGCTTGTGTTGCTCGCAGCGGCTCTGGAGCAAACT 1322
|||||
421 orleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLysLe 441
|||||
1323 TATCGGAACCGCAACTCTGTGTCTCTCTCGGAATACACTCTCTTCCATGGCTGCT 1382
|||||
441 uIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLeuLe 461
|||||
1383 AGGCTGTGTGCAACTGGATCTCGCGGAGTCTTGTCTAGCTCCGTCGCGCT 1442
|||||
461 uGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAlaLe 481
|||||
1443 GAATCCCGGCGAGCACCCGCTCTCGGGCGGCTTGGGGCTCTACCGTCCCTCTTCACTCT 1502
|||||

Db
481 uAenProAlaAspAspProSerArgGlyArgLeuGlyIleTyrArgProLeuLeuArgLe 501
QY
1503 GCGTTCGCGCGCACACACGCGGCGCACCTCTCTTTACGCGGTCTCCCGTATGTGCTTTC 1562
|||||
501 uProPheGlnProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValProSe 521
|||||
1563 TCATCTGCGCGGACCGGTGTCACCTTCGCTTCACCTCTGACGCTGCGATGGAGACCCGCTG 1622
|||||
521 rHisLeuProAspArgValIlePheAlaSerProLeuHisValAlaTrpArgProPro-- 540
|||||
1623 AACGCAACGCGAGGTCTTGCCTCAAGGTCTTATATAAGAGGACTCTTGGACTCTCAGCAATG 1682
|||||
540 ----- 540
QY
1683 TCAACGACGACCTTGAGGCATCTTCAAAGACTGTGTGTTTAAAGACTGGGAGGAGTTG 1742
|||||
540 ----- 540
QY
1743 GGGAGGAGATTAGGTTAAAGATTATGTACTAGGAGGCTGTAGGCNTAAATTTGCTGT 1802
|||||
540 ----- 540
QY
1803 TCACCAGCACCATGCAACTTTTCTCCTCTGCCTAATCATCTCATCTTCATGTCCTACTG 1862
|||||
540 ----- 540
QY
1863 TTCAAGCCTCAAGCTGTGCTGCTTGGGTGGCTTGGGACATGGACATGACCCGTATAAAG 1922
|||||
540 ----- 540
QY
1923 AATTTGGAGCATCTGCTGAGTTACTCTCTTTTGGCTTCTGACTTCTTCCGTCTATTTC 1982
|||||
540 ----- 540
QY
1983 GAGATCTCTCGACACCGCTCTGCTCTGTATCGGAGGCTTACAGTCTCCGGAACATT 2042
|||||
540 ----- 540
QY
2043 GTTGGCTTCACCATACAGCCTCAGGCAAGCTATTTTGTGTTGGGTGAGTTGATGAATC 2102
|||||
540 ----- 540
QY
2103 TGGCCACTGGGTGGGAGTAATTTGGAAGATCCAGCATCCAGGAATTAGTAGTCAGCT 2162
|||||
540 ----- 540
QY
2163 ATGTCAAAGTTAATATGGGCTTAAAACTCAGACAAATATTGTGTTTTCACATTTCTCTGTC 2222
|||||
540 ----- 540
QY
2223 TTACTTTTGGAAAGAAACTGTTCTTGAGTACTTGTGTATCTTTTGGAGTGTGATTCGCA 2282
|||||
540 ----- 540
QY
2283 CTCCTACCGCTTACAGACACCAATGCCCTTATCTTATCAACACTTCCGGAACACTACTG 2342
|||||
541 -----Met--ProLeuSerTyrGlnHisPheArgGlnLeu 552
|||||
2343 TTGTTAGCAGCAGCGAGGTCCCTCTAGAAAGAACTCCCTCGCTCGCAGACGGAAGGT 2402
|||||
553 LeuLeuAspGluAlaGlyProLeuGluGluLeuProArgLeuAlaAspGluAsp 572
|||||
2403 CTCATCGCGCGGTGCGAGAAGATCTCAATCTCGGAATCTCAAGTTAGTATCTCTGG 2462
|||||
573 LeuAsnArgArgValAlaGluAspLeuAsnLeuGlnLeuProAsnValSerIleProTrp 592
|||||
2463 ACTCTAAGGTGGAACTTACTGGGCTTATCTTCTACTGTACTGTCTCTTCTTAAATCCC 2522
|||||
593 ThrHisLysValGlyAsnPheThrGlyLeuTyrSerSerThrIleProValPheAsnPro 612
|||||
2523 GAGTGGCAAAATTCCTTCTCTTCTCACATTCATTTTCAAGAGGACATTATTAATAGATGT 2582
|||||
613 AsnTrpLysThrProSerPheProAspIleHisLeuHisGlnAspIleIleAsnLysCys 632
|||||

```
QY 2583 CAACAATATGTGGCCCTCTTACAGTAAATGAAAAAGAGATTAATAATTATGCGCT 2642
Db 633 GluInPheValGlyProLeuThrValAsnGluLysArgLeuAsnLeuValMetPro 652
QY 2643 GCTAGGTTTATCTAACCTTACTAATATTTGCCCTTAGACAAAGGCATTAACCGTAT 2702
Db 653 AlaArgPhePheProIleSerThrLysTyLeuProLeuGluLysGlyIleLysProTyr 672
QY 2703 TATCTGGAACATGACAGTAAATCACTTCTCAAACTAGGATTAATTACATCTCTGTGG 2762
Db 673 TyrProAspAsnValValAsnHisTyrPheGlnThrArgHisTyrLeuHisThrLeuTrp 692
QY 2763 AAGCTGGCATTTATATAGAGAGAACTACACGACGCGCTCATTTCTGGTGGTACCA 2822
Db 693 LysAlaGlyIleLeuTyrLysArgLysGluThrArgSerAlaSerPheCysGlySerPro 712
QY 2823 TATCTTGGGAACAAGAGCTACAGCATGGGAGGTGGTCTTCCAAACCTCGACAAGGCAT 2882
Db 713 TyrSerTrpGluGlnGluLeuHisGlyAlaPheLeu---AspGlyProSerArgMet 731
QY 2883 GGGAGACAACTTGTCTGTTCCCAATCTCTGGGATTTCTTCCGATCACCAGTTGGACCC 2942
Db 732 GlyGluGluTyrPheHisHisGlnSerSerGlyIlePheSerArgProProValGlySer 751
QY 2943 TGCCTCGGAGCCAACTCAACAATCCAGATTGGGATTTCAACCCCAACAGGATCACTG 3002
Db 752 SerIleGlnSerLysHisGlnLysSerArgLysGlyProGlnSerGlnGlnArgProLeu 771
QY 3003 GCCAGAGCCAACTCAGGTAGGAGTGGGAGCATTCGGGCGGAGTTACCCGCCACACACGG 3062
Db 772 AspArgSerGlnGlnGlyArgSerGlySerIleArgLalaTrpValHisSerProThrArg 791
QY 3063 CGGTCTTTTGGGGGGGAGCCCTCAGGCTCAGGCGATATTGACAAAGTGGCAGCAGCAC 3122
Db 792 ArgProPheGlyValGluProSerGlySerArgHisAlaLysAsnIleAlaSerArgSer 811
QY 3123 TCCTCTGCTCCCAATCGGAGTGGGAGTGGGAGCAGCTACTCCCATCTCTCCACCTCT 3182
Db 812 AlaSerCysLeuHisGlnSerAlaValArgLysAlaLalaTyrProAsnHisSerThrPhe 831
QY 3183 AAGAGACAGTCACTCCTCAGGCGCAGCAGTGGAA 3215
Db 832 GluArgHisSerSerSerGlyHisAlaValGlu 842

RESULT 4
ID Q7TDR5_HPBVO PRELIMINARY; PRT; 843 AA.
AC Q7TDR5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polymerase protein.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Liu C.-J., Chen P.-J., Lai M.-Y., Kao J.-H., Chen D.-S.;
RL Submitted (Oct-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY167095; AAO41311.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006278; F:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001462; DNAPol_viral_C.
DR InterPro; IPR000201; DNAPol_viral_N.
DR InterPro; IPR000477; RVTse.
```

```
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00078; RVT_1; 2.
DR PRODom; PD000814; DNAPol_viral_C; 1.
SQ SEQUENCE 843 AA; 94471 MW; A40B9C18D84B4EA5 CRC64;

Alignment Scores:
Pred. No.: 6,32e-202 Length: 843
Score: 2824.00 Matches: 530
Percent Similarity: 98.52% Conservative: 3
Best Local Similarity: 97.97% Mismatches: 7
Query Match: 47.81% Indels: 2
DB: Gaps: 0

US-10-761-006a-1 (1-3215) x Q7TDR5_HPBVO (1-843)
QY 1 CTCACAACATTTCCACAAGCTCTGCTAGTATCCAGGCTGAGGGCTATATTTCTCTGC 60
Db 304 LeuHisAsnIleProSerSerThrArgSerGlnSerGluGlyProIlePheSerCys 323
QY 61 TGGTGGCTCCAGTTCCGGAACAGTAAACCTGTTCCGACTACTGCTCTCCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgHisSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
QY 121 AATCTTCTCGAGGACTGGGACCTCGACCGAACAATGGAGAACACACATCAGGATTCCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGAGCCCTCTCGTGTACAGGCGGGGTTTTCTCGTTCACAGATCTCTCAATATACC 240
Db 364 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 GCAGAGCTAGACTCG-GTGGACTTCTCTCAATTTCTAGGGGAGACCCACGCTGTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgLysSerThrHisVal-Se 403
QY 300 CTGGCCAAAATTCGAGTCCCACTCCCAATCACTCACCAACCTCTGCTCTCCAATTT 359
Db 403 rTTPProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
QY 360 GTCTGCTATCGTGGATGTGTCTGGGGGTTTTATCATATTCTCTCTCATCTCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
QY 420 TATGCTCATCTCTCTGTTCTCTGCTGACTACCAAGGTATGTTGCCCTTTGCTCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
QY 480 TACTTCCAGGAACATCAACCAACAGCAGCGGGCCATGCAAGACCTGCAAGCTCTGCTC 539
Db 463 rThrSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSe 483
QY 540 AAGGAACTCTACGTTTCCCTTCTGTTGCTGTACAAAACCTTCGACGGGAACTGCACCT 599
Db 483 rArgAsnLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
QY 600 GTATTCATCCCATCATCTCTGGGCTTTCCGAAGATTCCTATGGAGTGGGCTCAGTCC 659
Db 503 uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
QY 660 GTTCTCTCTGCTCAGTTTACTAGTGCATTTGTTTCTAGTGTGTTCTGAGGGCTTTCC 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHis 543
QY 720 CTGTTTGGCTTTCAGTTATATGAGATGATGTGGTATTTGGGGCGGAGTCTGTACAACATCT 779
Db 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTTAAACCC 839
Db 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
QY 840 TAATAAAACCAACGTTGGGGCTACTCCCTTAACTTCATGGGATATGTAATTGGAAGTTG 899
```

|||||
583 oAnuLyThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
QY GGGTACTTTTACCGCAGGAACATATTGCTATAAATCAAGCAATGTTTTCGAAAACTGCC 959
|||||
603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
QY TGTAAATAGACCTATTGATTGGAAAGATGTCAAAGATTGGGGCTTTTGGGCTTTGCG 1019
623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY TGCCCTCTTTACACAATAGGCTATCTGCTCTGATGCGCTTTATATGCGATGTATACAATC 1079
643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
QY TAAGCAGGCTTTCACTTTCTCGCAACTTACAGGCTTTCTGTGTAAACAATATCTGAA 1139
663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 683
QY CTTTACCCGCTTGGCGGCAAGGTCGGTCTCTGCCAAGTGTTCGCTGACGCAACCC 1199
683 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
QY CACTGTATGGGCTTGGCCATAGCCATCAGCGCATGGCTGGAACTTTCTGGCTCTCTCT 1259
703 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 723
QY GCGATCCCATCTCGGAACTCTAGCAGCTTGTGCTCGAGCGGCTCTGGAGCAAA 1319
723 uProIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLys 743
QY ACTTATCGGAACCGCAACTCTGTTGCTCTCTCGGAATATACACTCTCTTCATGGCT 1379
743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
QY GCTAGGCTGTCGCCCACTGGATCTGCGCGGAGCTCTGTTGCTCTAGCTCCGTCGGC 1439
763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
QY GCTGAATCCCGCGAGACACCGCTCTCGGGCCCTTTGGGGCTCTACCGTCCCTCTTCA 1499
783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
QY TCTGCGGTTCCGGCCGACACACGGGGCGACCTCTCTTTACGGGTCTCCCGTATGTGCC 1559
803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
QY TTCTCATCTCGCGGACCGTGTGCACCTTCACCTCTGCAGCTGCATGGAGACCACC 1619
823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaIleArgProPr 843
1620 G 1620
843 o 843

RESULT 5

Q4FD83 HPBV0 PRELIMINARY; PRT; 843 AA.
ID Q4FD83;
AC Q4FD83;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Polymerase.
DE Name=P;
GN Hepatitis B virus.
OS Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthonepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HK343;
RX PubMed=15897987; DOI=10.1086/430324;
RA Chan H.L., Tsui S.K., Tse C.H., Ng E.Y., Au T.C., Yuen L.,

RA Bartholomeusz A., Leung K.S., Lee K.H., Locarnini S., Sung J.J.;
RT "Epidemiological and virological characteristics of 2 subgroups of
RT hepatitis B virus genotype C";
RL J. Infect. Dis. 191:2022-2032(2005).
DR EMBL; D0089793; AA05280.1; -; Genomic DNA.
SQ SEQUENCE 843 AA; 94494 MW; BB50233DFEBA4E29 CRC64;

Alignment Scores:
Pred. No.: 8,92e-202 Length: 843
Score: 2822.00 Matches: 530
Percent Similarity: 98.52% Conservative: 3
Best Local Similarity: 97.97% Mismatches: 7
Query Match: 47.77% Indels: 2
DB: Gaps: 0

US-10-761-006A-1 (1-3215) x Q4FD83_HPBV0 (1-843)

QY 1 CTCACAACATTTCCACCAAGCTCTGTAGATCCCAAGGTGAGGGCGCTATATTTCTCTGC 60
Db |||||
304 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 323
QY 61 TGGTGGCTCCAGTTCCGGAACAGTAAACCTCTTCCGACTACTGCTCTCCCATATCGTC 120
Db |||||
324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
QY 121 AATCTTCTCGAGACTGGGGACCTGACCGAACATGAGAGAACACACATCAGGATTCCT 180
Db |||||
344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGGACCCCTGCTCGTGTACAGCGGGTTTTCTCGTTGACAAGAATCCTCACATAACC 240
Db |||||
364 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 GCAGAGTCTAGACTCTG-GTGACTTCTCTCAATTTCTAGGGGGAGCACCCACGTGTTCT 299
Db |||||
384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
QY 300 CTGGGCAAAATTCGAGTCCCCCAACTCAATCACTACCAACCTCTTGTCTCTCCAAATT 359
Db |||||
403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
QY 360 GTCCTGGCTATCGCTGGATGCTGCGGCGTTTTATCATATTTCTCTTTCATCTCTGCTGC 419
Db |||||
423 userTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
QY 420 TATGCTCATCTCTTGTGTTGTTCTTCTGAGCTACCAAGGTATGTTGCCGTTGTCCTC 479
Db |||||
443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
QY 480 TACTTCCAGGAACATCAACCAACAGCAGCGGGCCATGCAAGACCTGCACGACTCTCTGCTC 539
Db |||||
463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
QY 540 AAGGAACTCTACGTTTCCCTCTGTTGCTGTACAAAACCTTCGGACGGAACCTGCACATT 599
Db |||||
483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
QY 600 GTATTCCCATCCCATCATCTCTGGGCTTTCGCAAGATTCCTATGGGAGTGGGCTCAGTCC 659
Db |||||
503 uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
QY 660 GTTCTCTCGGCTCAGTTTACTAGTGCCATTGTTGTCAGTGGTTCGTAGGGCTTTCCCCCA 719
Db |||||
523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
QY 720 CTGTTTGGCTTTTCAGTTATATGATGATGTTGTTGGGGCGAAGCTCTGTCAACATCT 779
Db |||||
543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
Db |||||
563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 593

QY 840 TAATAAACAACAGTTGGGGCTACTCCCTTAATCTCATGGATATGTAATGGAAGTTG 899
Db 583 oAenLysThrLysArgTrpGlyTrpSerLeuAnpPheMetGlyTrpValIleGlySerTr 603
QY 900 GGGTACTTTTACCAGGAAACATATGTTACTAAACCTCAAGCAATGTTTTCGAAAACTGCC 959
Db 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAAATAGACCTATTGATTGGAAAGTATGTCACAAAGATTGTCGGCTTTTGGCTTTC 1019
Db 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGGCCCTTTTACAAATGTCCTATCTGCTGATGCCCTTTATATGATCATGTATACAATC 1079
Db 643 aAlaProPheThrGlnCysGlyTrpProAlaLeuMetProLeuTrpAlaCysIleGlnSe 663
QY 1080 TAAAGCAGGCTTTTCACTTTCTCGCAACTTCAAGGCTTTTCTGTGTAAACAATATCTGAA 1139
Db 663 rLysGlnAlaPheThrPheSerProThrTrpLysAlaPheLeuLysGlnTrpLeuAs 683
QY 1140 CTTTACCCGCTGTCGGGCAACGGTCCGCTCTCTGCCAAGTGTGTCGACCAACCCC 1199
Db 683 nLeuTrpProValAlaArgGlnArgSerGlyLeuLysGlnValPheAlaAspAlaThrPr 703
QY 1200 CACTGGATGGGCTTGGCCATAGCCATCAGCGATCGCTGGAACTTTCTGCTGCTCTCT 1259
Db 703 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 723
QY 1260 GCCATCCATCTCGGAACTCTCTAGCAGCTGTTTGTCTGCGACCGGCTCGAGCAAA 1319
Db 723 uProIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLy 743
QY 1320 ACTTATCGGAACCGCAACTCTGTGTCTCTCTCGGAAATACACCTCTCTTCCATGCT 1379
Db 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTrpThrSerPheProTrpLe 763
QY 1380 GCTAGGTGTGCTGCCAAGTCTGCTGCGGAGCTCTCTGCTGCTGCTGCTGCTGCTG 1439
Db 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTrpValProSerAl 783
QY 1440 GCTCAATCCCGGACGACCGCTCTCGGGGCGGCTTGTGGGCTCTACCGTCTCCCTTCTCA 1499
Db 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuLysLeuTrpArgProLeuLeuHi 803
QY 1500 TCTCCGCTTCCGGCCGACCGGCGGCGACCTCTCTTTTACCGGCTCTCCCGTATGTCGC 1559
Db 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTrpAlaValSerProSerValPr 823
QY 1560 TTCTCATCTGCGGACCGTGTGCACTTCCGCTTCCACCTGCGACGTCGATGAGACCAACC 1619
Db 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
QY 1620 G 1620
Db 843 o 843

RESULT 6
ID Q69026_HPBVO PRELIMINARY; PRT; 540 AA.
AC Q69026;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE P gene product (Fragment).
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AYT;
RX MEDLINE=87059755; PubMed=3783127;
RA Okamoto H., Imai M., Shimozaki M., Hoshi Y., Iizuka H., Gotanda T.,

RA Tsuda F., Miyakawa Y., Mayumi M.;
RT "Nucleotide Sequence of a Cloned Hepatitis B Virus Genome, Subtype
RT avr:Comparison with Genomes of the Other Three Subtypes.";
RL J. Gen. Virol. 67:2305-2314(1986).
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; X04615; CAA28286.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001462; DNAPol_viral_C.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00078; RVT_1; 1.
DR ProDom; PD000814; DNAPol_viral_C; 1.
DR PROSITE; PS00878; RT_POL; 1.
FT NON_TER 1
SQ SEQUENCE 540 AA; 60365 MW; 80446CA7C80F998A CRC64;

Alignment Scores:
Pred. No.: 6 02e-201 Length: 540
Score: 2810.00 Matches: 527
Percent Similarity: 98.34% Conservative: 5
Best Local Similarity: 97.41% Mismatches: 8
Query Match: 47.57% Indels: 2
DB: Gaps: 0

US-10-761-006A-1 (1-3215) x Q69026_HPBVO (1-540)
QY 1 CTCACAAATTCACCAAGCTCTGTAGATCCAGGGTCCAGGGCCATATATTTCTCTGC 60
Db 1 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 20
QY 61 TGTGTGCTCAGTTCGGAAACAGTAAACCTGTTCGACTACTCCCTCTCCCATATGTC 120
Db 21 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTrpCysLeuThrHisIleVal 40
QY 121 AATCTTCTCAGGACTGGGACCTCGACCGAATCATGAGAACACACATCAGGATTCCT 180
Db 41 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 60
QY 181 AGACCCCTCTCTGTGTACAGCGGGGTTTTCTCGTTGACAAAGATCTCTCAATACC 240
Db 61 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 80
QY 241 GCAGAGTCTAGACTGTG-GTGGACTTCTCTCAATTTCTAGGGGAGGACCCAGTGTTC 299
Db 81 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 100
QY 300 CTGGCCAAATTCGACGTCCCACTCCCACTCACTCACTCACTCACTCTTGTCTCTCAATT 359
Db 100 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 120
QY 360 GTCCTGGCTATCGTGGATGTCTGCGGGCGGTTTTTATCATATTTCTCTCTCTCTCTGC 419
Db 120 uSerTrpLeuSerLeuAspValSerAlaAlaPheTrpHisIleProLeuHisProAlaAl 140
QY 420 TATGCTCATCTTCTGTGGTCTTCTGACTACCAAGGTATGTTGCCCGTTGCTCTC 479
Db 140 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTrpValAlaArgLeuSerSe 160
QY 480 TACTTCCAGGACATCAACACCGACGCGGGCCCATGCAAGCTTGCAGCAGTCTCTCTCTC 539
Db 160 rThrSerArgAsnIleAsnTrpGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 180
QY 540 AAGGAACTCTACGTTTCCCTCTGTTGCTGTCTGACAAACCTTCGACGAGAACTGCAC 599
Db 180 rArgAsnLeuTrpValSerLeuLeuLeuLeuLeuTrpLysThrPheGlyArgLysLeuHis 200
QY 600 GTATTTCATCCATCATCTCTGGGCTTTCGAAAGATTCCTATGGAGTGGGCTCAGTCC 659

Db 200 uTyrSerHisProIleLeuGlyPheAArgLysIleProMetGlyValGlyLeuSerPr 220
QY 660 GTTTCCTCGGCTCAGTTACTAGTGCATTTCTTTCAGTGGTTCGTAGGCTTTCCCA 719
Db 220 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 240
QY 720 CTGTTTGGCTTTCAGTTATATGATGATGTGGTATTGGGGCCGAAGTCTGTACAAATCT 779
Db 240 sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe 260
QY 780 TGAGTCCCTTTTACCTTATTACCAATTTCTTTCTTTGGGTATACATTTAAACCC 839
Db 260 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 280
QY 840 TATAAACAACCGTTGGGGCTACTCCCTTACTTATCGGGATATGTAATTCGAAGTTG 899
Db 280 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 300
QY 900 GGTGACTTTACCCGAGCAACATATTGTACTAAACTCAAGCAATGTTTTCGAAAACCTGCC 959
Db 300 pGlyThrLeuProGlnGluHisIleValGlnLysLeuLysGlnCysPheArgLysLeuPr 320
QY 960 TGTAAATAGACCTATTGATTGGAAGTATGTCAAAGAAATTGTGGTCTTTTGGGCTTTGC 1019
Db 320 oValAsnArgProIleAspTrpLysValCysGlnArgIleValIleGlyLeuLeuGlyPheAl 340
QY 1020 TGCCCTTTTACACATGGGCTATCTGCTGTGATGCCCTTATATGATGATGATATACAATC 1079
Db 340 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 360
QY 1080 TAAGCAGGCTTTCACATTTCTCGCAACTTACAAAGGCTTCTGTGTAAACAATATCTGAA 1139
Db 360 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 380
QY 1140 CCTTTACCCCGTTGCCGCAACGGTCCGCTCTCTGCCAAGTGTGTTGCTGACGCAACCC 1199
Db 380 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 400
QY 1200 CACTGATGGGCTTGGCCATAGGCATCAGCCATGGCTGGAACTTCTTGCTGCTCTCT 1259
Db 400 oThrGlyTrpGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe 420
QY 1260 GCGATCCCATCTGCGGAACCTCTAGCAGCTTGTGTTTGTCTGCGAGCGGTCTGGAGCAA 1319
Db 420 uProIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLyl 440
QY 1320 ACTTATCGGAACGCAACTCTGTTGCTCTCTCGGAAATACACCTCTCTTCCATGGCT 1379
Db 440 sLeuIleGlyThrAspAsnSerValValLeuSerAArgLysTyrThrSerPheProTrpLe 460
QY 1380 GCTAGGTTGTGCGCAACTGGATCTCTGCGCGGAGCGTCTTGTCTACGTCCCGTCGCGC 1439
Db 460 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 480
QY 1440 GCTGAATCCCGGAGCAGCCGCTCGGGCGCTTTGGGGCTCTACCGTCCCTTCTTCA 1499
Db 480 aLeuAsnProAlaAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 500
QY 1500 TCTGCGGTTCCCGCCACACCGGGCGCACCTCTCTTTTACGCGGTCTCCCGCTATGTGC 1559
Db 500 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 520
QY 1560 TTCTCATCTCGCGGACCGTGTGACCTTGTTCACCTCTGCGAGTGCATGGAGACACCC 1619
Db 520 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 540
QY 1620 G 1620
Db 540 o 540

RESULT 7
Q9QN49 HPBV0

ID Q9QN49_HPBV0 PRELIMINARY; PRT; 843 AA.

AC Q9QN49; (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 26, Last annotation update)
DE Polymerase.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=genotype C;
RX MEDLINE=20528425; PubMed=11074476;
RX DOI=10.1002/1096-9071(200012)62:4<4471::AID-JMV12>3.3.CO;2-O;
RA Rokuhara A., Tanaka E., Yagi S., Mizokami M., Hashikura Y.,
RA Kawasaki S., Kiyosawa K.;
RT "De novo infection of hepatitis B virus in patients with orthotopic
RT liver transplantation: analysis by determining complete sequence of
RT the genome.";
RL J. Med. Virol. 62:471-478 (2000).
DR EMBL; AB026815; BAA85065.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001462; DNAPol_viral_C.
DR InterPro; IPR000201; DNAPol_viral_N.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00078; RVT 1; 2.
DR ProDom; PD000814; DNAPol_viral_C; 1.
SQ SEQUENCE 843 AA; 94462 MW; 9D741241062DEAAE CRC64;

Alignment Scores:
Pred. No.: 9,92e-201 Length: 843
Score: 2808.00 Matches: 527
Percent Similarity: 98.52% Conservative: 6
Best Local Similarity: 97.41% Mismatches: 7
Query Match: 47.54% Indels: 2
DB: Gaps: 0

US-10-761-006A-1 (1-3215) x Q9QN49_HPBV0 (1-843)
QY 1 CTCCCAACATTCACCAAGCTCTGTAGATCCAGGGTGGGGCGCTATATTTCTCTGC 60
Db 304 LeuHisHisIleProProSerSerAlaArgSerGlnSerGlyProIlePheSerCys 323
QY 61 TGTGTGCTCCAGTTCGGGAACAGTAAACCTGTTCGACTACTGCTCTCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
QY 121 AATCTTCTCGAGCATGGGGACCTTCACCAACATGGAGAACACAACTCAGGATTCCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGGACCCCTGCTCGTGTATACAGCGGGGTTTTCTCGTTGACAAAGTAATCCTCACATACC 240
Db 364 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 CGAGACTCTAGACTCTG-GTGGACTTCTCTCAATTTCTTAGGGGGAGCACCCACGCTGTTTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
QY 300 CTGGGCAAAATTCGAGTCCCCCAACTCAATCACTCACCACCTTCTCTCTCCCAATTT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
QY 360 GTCTGGCTATCGCTGGATGTCTCTGGCGGCTTTTATCATATTTCTCTTCATCTCTCTCTGC 419

Db 423 userTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeProLeuHisProAlaAl 443
 Qy 420 TATGCCCTCATCTCTCTGTTGTTCTTCTGAGTACCAAGTATGTTGCCCGTTGCTC 479
 Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
 Qy 480 TACTTCCAGAAATCAACACACGACGCGGCGCATGCAAGACTGCGACGACTCTGCTC 539
 Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
 Qy 540 AAGGAAACTCTACGTTTCCCTCTCTGTTGCTGTACAAACCTTTCGACGCGAACTGCACTT 599
 Db 483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
 Qy 600 GTATTCCTCCATCCATACCTCGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
 Db 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
 Qy 660 GTTTCCTCTGCTCAGTTACTAGTGCATTTGTTGAGTGTTCAGTGGTTCAGTGGCTTCC 719
 Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHI 543
 Qy 720 CTGTTTCTGCTTTCAGTTATATGATGATGATGATGATGATGATGATGATGATGATGAT 779
 Db 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
 Qy 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTG 839
 Db 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
 Qy 840 TAATAAACCAACGTTGGGCTACTCCCTTAACTCATGATGGATATGTAATGCAAGTTG 899
 Db 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
 Qy 900 GGGTACTTTACGCGAGGAACATATTGTAATAAATCAAGCAATGTTTTCGAAACTGCC 959
 Db 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
 Qy 960 TGTAAATAGACTATTGATGGAAAGTATGTCAAGAAATGTTGTTGTTGTTGTTGTTGTTG 1019
 Db 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
 Qy 1020 TGCCCTTTTACATATGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1079
 Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
 Qy 1080 TAACGAGGCTTTCATTTCTCGCAACTTACAGGCTTTCTGTGTAACAATATCTGAA 1139
 Db 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi 683
 Qy 1140 CTTTACCCCGTGTCCCGCAACGCTCCGCTCTCTGCCAAGTGTGTTGCTGACGCAACCCC 1199
 Db 683 sLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
 Qy 1200 CACTGATGGGCTTGGCCATAGCCGATAGCCGATAGCCGATAGCCGATAGCCGATAGCCG 1259
 Db 703 oThrGlyTrpGlyLeuAlaIleGlyHisArgGlnMetArgGlyThrPheValAlaProLe 723
 Qy 1260 GCCGATCCATCTCGGAACTCTAGCAGCTGTTTGTCTGCGACCGCTCTGGAGCAAA 1319
 Db 723 uProfileHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLys 743
 Qy 1320 ACTTATCGAACCGCAACTCTGTTGCTCTCTCTCGGAAATACACCTCTTTTCCATGCT 1379
 Db 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTiple 763
 Qy 1380 GCTAGGGTGTGTCCAACTGATCCTGCGGAGCTCTTTGCTGCTGCTGCTGCTGCTGCTG 1439
 Db 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
 Qy 1440 GCTGAATCCGCGGACGACCTCTCGGGCGGTTTGGGCTCTACCGTCCCTCTCTTCA 1499

Db 783 aLeuAsnProAlaAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
 Qy 1500 TCTCGCGTTCGCGCCACACACGCGCGCCTCTCTTTACGCGTCTCCCGTATGTGCC 1559
 Db 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
 Qy 1560 TTCTCATCTGCGGACCGTGTGCACTTCGTTTCACTTCGCTGCGACGTCGATGGAGACC 1619
 Db 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
 Qy 1620 G 1620
 Db 843 o 843

RESULT 8

DPOL_HPBVR
 ID DPOL_HPBVR STANDARD; PRT; 843 AA.
 AC P03157;
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE P protein [includes: DNA-directed DNA polymerase (EC 2.7.7.7); RNA-
 directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
 GN Name:P;
 OS Hepatitis B virus (subtype adr).
 OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
 OC Orthohepadnavirus.
 OX NCBI_TaxID=106820;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
 RA MEDLINE=83168919; PubMed=6300776;
 RA Ono Y., Onda H., Sasada R., Igarashi K., Sugino Y., Nishioka K.;
 "The complete nucleotide sequences of the cloned hepatitis B virus
 DNA; subtype adr and adr."
 RT Nucleic Acids Res. 11:1747-1757(1983).
 RL Nucleic Acids Res. 11:1747-1757(1983).
 CC -1- CATALYTIC ACTIVITY: Deoxynucleoside triphosphate + DNA (n) =
 diphosphate + DNA(n+1).
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 phosphomonoester.
 CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use as long as its content is in no way modified and this statement is not
 removed.
 CC -----
 CC EMBL; V00867; -; NOT_ANNOTATED_CDS; Genomic_DNA.
 DR PIR; A00704; JDVLR.
 DR InterPro; IPR001462; DNAPol_viral_C.
 DR InterPro; IPR000201; DNAPol_viral_N.
 DR InterPro; IPR000477; RVSe.
 DR Pfam; PF00336; DNA_pol_viral_C; 1.
 DR Pfam; PF00242; DNA_pol_viral_N; 1.
 DR Pfam; PF00078; RVT_1; 1.
 DR ProDom; PD000814; DNAPol_viral_C; 1.
 DR ProSITE; PS00878; RT_POL_1.
 DR DNA replication; DNA-binding; DNA-directed DNA polymerase;
 Endonuclease; Hydrolase; Multifunctional enzyme; Nuclease;
 KW Nucleotidyltransferase; RNA-directed DNA polymerase; Transferrase.
 FT DOMAIN 357 600 Reverse transcriptase.
 SQ SEQUENCE 843 AA; 94400 MW; A6B2D490839C4E8B CRC64;
 Alignment Scores:
 Pred. No.: 1.18e-200 Length: 843
 Score: 2807.00 Matches: 528
 Percent Similarity: 98.33% Conservative: 3
 Best Local Similarity: 97.78% Mismatches: 8
 Query Match: 47.52% Indels: 2
 DB: 1 Gaps: 0

US-10-761-006a-1 (1-3215) x DPOL_HPBVR (1-843)

QY 4 CACAACTCCACCAAGCTCTGCTAGATCCAGGCTGAGGGCGCTATATTTCTCTGCTGG 63
 Db |||||
 305 HisAsnIleProSerSerAlaAargSerGlnSerGlnGlyProIlePheSerCysTrp 324
 QY 64 TGGCTCCAGTTCGGGAACAGTAACCCCTGTTTCGGACTACTGCTCTCCCATATCGTCAAT 123
 Db |||||
 325 TrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAsn 344
 QY 124 CTTCTCAGAGCTGGGGACCCCTGCACCGAAATCGAGAAACAACTCAGGATTCCTAGG 193
 Db |||||
 345 LeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArg 364
 QY 184 ACCCTGCTCGTGTACAGCGGGGTTTTCTCGTTGACAGAAGATCCTCACATACCGCA 243
 Db |||||
 365 ThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrThr 384
 QY 244 GAGTCTAGACTCTG-GTGGACTCTCTCAATTTTCTAGGGGAGACCCACGCTGTTCCTG 302
 Db |||||
 385 GluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-SerTr 404
 QY 303 GCCAAATTCGAGTCCCGAACCTCCAACTCACTCAACACCTCTGTGCTCGAATTTGTC 362
 Db |||||
 404 pProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLeuSe 424
 QY 363 CTGGCTATCGCTGATGTGTCGGCGGTTTTATCATATTTCTTCATCCTGCTGCTAT 422
 Db |||||
 424 rTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAlaMe 444
 QY 423 GCCTCATCTTTCTGTTGTTCTCTCGACTACCAAGGTATGTTGCCGTTTGTCTCTCTAC 482
 Db |||||
 444 tProHisLeuLeuValValGlySerSerGlyLeuProAArgTyrValAlaArgLeuSerSerTh 464
 QY 483 TTCAGGAACATCAACACGAGCCGCGCCATGCAAGACTGCAAGCTCTGCTGCTCAAG 542
 Db |||||
 464 rSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSerAr 484
 QY 543 GAAACTCTAGTTTCCTCTGTTGTGTACAAAACCTTCGGACGNAACGTCACATTGTA 602
 Db |||||
 484 gAsnLeuTyrValSerLeuLeuLeuLeuTyrHisThrPheGlyArgLysLeuHisLeuTyr 504
 QY 603 TTCCCATCCCATCATCTCTGGGCTTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCCGTT 662
 Db |||||
 504 rSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPh 524
 QY 663 TCTCTCGCTCAGTTTACGTACGTCCTTGTTCAGTGGTGTGAGGGCTTCCCCCAGTCG 722
 Db |||||
 524 eLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHisCy 544
 QY 723 TTTGGCTTTTCAGTTATATGATGATGTTGTTATGGGGCGAAGTCTGTACACATCTTGA 782
 Db |||||
 544 sLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGl 564
 QY 783 GTCCCTTTTACCTTATACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCCCTAA 842
 Db |||||
 564 uSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAs 584
 QY 843 TAAACCAACGTTGGGCTACTCTCTTAATTCATGAGATATGTAATTTGGAAGTTGGGG 902
 Db |||||
 584 nLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrpGl 604
 QY 903 TACTTTACCGCAGGACATATTTGTTACTTAAACCTCAAGCAATGTTTTCGAAACCTGCTCT 962
 Db |||||
 604 yThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuProVa 624
 QY 963 AAATAGACCTATTGATTGGAAGATGTCAAGAAATTTGGGGTCTTTTGGGCTTTCTGCTG 1022
 Db |||||
 624 lAsnSerProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaI 644
 QY 1023 CCTTTTACAAATGCGGTATCCTCGCTGTATGCTTTTATATGATATGATACAACTAA 1082
 Db |||||
 644 aProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerLy 664
 QY 1083 GCAGGCTTTTCCTCGCCAACTTACAAGGCTTTCTGTGTAAACAATAATCTGAACCT 1142

Db 664 sGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsnLe 684
 QY 1143 TTATCCCGGTTGCCCGCAACGGTCCGGTCTCTGCCAAGTGTTTTGTGACGCAACCCCCAC 1202
 Db 684 uTyrProValAlaAargGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProTh 704
 QY 1203 TGGATGGGGCTTGGCCATAGGCCATCAGCGCATCGGCTGGAACCTTTCTGGCTCTCTGCC 1262
 Db 704 rGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLeuPr 724
 QY 1263 GATCCATCTATCGGGAACCTCTAGCAGCTGTTGTTGCTCGCAGCCGCTCTGGAGCAAAACT 1322
 Db 724 olleHisThrAlaGluLeuLeuAlaAlaCysPheAlaAargSerArgSerGlyAlaLysLe 744
 QY 1323 TATCGGAACCGCAACTCTGTTGCTCTCTCGGAATAACACTCTCTTCCATGGCTGCT 1382
 Db 744 uIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLeuLe 764
 QY 1383 AGGTTGCTGTCGAACCTGATCCTGCGCGGAGCGTCTTGTGCTACGTCCTCGTTCGGGCT 1442
 Db 764 uGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAlaLe 784
 QY 1443 GAATCCCGCGGACGACCCGCTCTCGGGCGGCTTTGGGGCTCTACCGTCCCTTCTCATCT 1502
 Db 784 uAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuLeuLe 804
 QY 1503 GCGGTTCCGCGCACACACGGGGCGACCTCTCTTTACGGGGTCTCCCGTATGTGCTTC 1562
 Db 804 uProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValProSe 824
 QY 1563 TCATCTGCGGACCGGTGTCACCTTCGCTTCACCTTCGACGTCGATGAGAGACCGG 1620
 Db 824 rHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 843

RESULT 9
 Q97975_HPBVO PRELIMINARY; PRT; 540 AA.
 AC Q97975;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE ORF (Fragment).
 OS Hepatitis B virus.
 OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
 OC Orthohepadnavirus.
 OX NCBI_TaxID=10407;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=95294549; PubMed=7775946;
 RA Uchida T., Gotoh K., Shikata T.;
 RT "Complete nucleotide sequences and the characteristics of two
 RT hepatitis B virus mutants causing serologically negative acute or
 RT chronic hepatitis B.";
 RL J. Med. Virol. 45:247-252 (1995).
 CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
 DR EMBL; D16666; BAA04073.1; -; Genomic DNA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0004523; F:ribonuclease H activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
 DR InterPro; IPR001462; DNAPol_viral_C.
 DR InterPro; IPR000477; RVase.
 DR Pfam; PF00336; DNA_pol_viral_C; 1.
 DR Pfam; PF00078; RVF_1; 1.
 DR ProDom; PD000814; DNAPol_viral_C; 1.
 DR PROSITE; PS50878; RT_POL; 1.
 FT NON_TER
 SQ SEQUENCE 540 AA; 60346 MW; 7BBE285359DD39C5 CRC64;

Alignment Scores:

Pred. No.:	1.2e-200	Length:	540
Score:	2806.00	Matches:	526
Percent Similarity:	98.34%	Conservative:	6
Best Local Similarity:	97.23%	Mismatches:	8
Query Match:	47.50%	Indels:	2
DB:	2	Gaps:	0

US-10-761-006A-1 (1-3215) x Q97975_HPBVO (1-540)

Qy	1	CTCCACAACTTCCACCAAGCTCTGCTAGATCCAGGGTGGGGCCCTATATTTCTCTGC	60
Db	1	LeuHisAsnLeuProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys	20
Qy	61	TGTTGGCTCAGTTCGGAACAGTAACCTGTTCCGACTACTCCCTCCCATATCGTC	120
Db	21	TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal	40
Qy	121	AATCTTCTCGAGGACTGGGACCTCGACCGAACATGGAGAACACAACTCAGGATTCCT	180
Db	41	AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro	60
Qy	181	AGGACCCCTGCTGTTTACAGCGGGGTTTTCTCGTTGACAAAGATCTCACAATACC	240
Db	61	ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr	80
Qy	241	GCAGACTAGACTCTG- GTGGACTCTCTCAATTTTCTAGGGGAGACCCAGCTGTC	299
Db	81	ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal- Se	100
Qy	300	CTGGCCAAATTCGACGTCCCACTCCCAATCACTACCAACCTCTTGCTCCCAATTT	359
Db	100	rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerAsnLe	120
Qy	360	GTCCTGGCTATCGCTGGATGTCTGGGGCGTTTTATCATATTCCTCTTCATCTGCTGC	419
Db	120	uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl	140
Qy	420	TATGCTCATCTCTGTTGGTCTCTGGAATCAAGGATGTTGGCGGTTTGCTCTC	479
Db	140	aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe	160
Qy	480	TACTTCCAGGAACATCAACACCGACGCGGCATGCAAGCTCGACGACTCTGCTC	539
Db	160	rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe	180
Qy	540	AAGGAACTCTACGTTTCCCTCTGTTGCTGTACAAAACCTTCGGACGGAAACTGCATT	599
Db	180	rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe	200
Qy	600	GTATTCCTCATCCATCATCTGCTGGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC	659
Db	200	uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr	220
Qy	660	GTTTCTCTGCTCAGTTTACTAGTGCATTTGTTTCACTGTTTCTAGGCTTTTCCCCCA	719
Db	220	oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi	240
Qy	720	CTGTTTGGCTTTCAGTATATGATGATGTGTGTTTGGGGCGCAAGTCTGTACACATCT	779
Db	240	sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe	260
Qy	780	TGAGTCCCTTTTACTCTATPACAAATTTCTTTTGTCTTTGGGTATACATTTAAACCC	839
Db	260	uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr	280
Qy	840	TAATAAACCAACCTTGGGCTACTCCCTTAACCTTCATGGATATGTAATTTGGAAGTTG	899
Db	280	oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr	300
Qy	900	GGGTACTTTTACCGCAGGAACATATTGTACTAAACCTCAAGCAATGTTTTCCGAAAACCTGCC	959
Db	300	pGlyThrLeuProGlnGluHisIleValLeuLeuLysLeuLysGlnCysPheArgLysLeuPr	320

Qy	960	TGTAATAGACCTATTGATTGGAAGATATGTCAAGAAATTTGGGTCTTTTGGGCTTTGC	1019
Db	320	oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl	340
Qy	1020	TGCCCCCTTTTACAAATGTGGCTATCTGCTTGTATGCTTATATATGATGATATACATC	1079
Db	340	aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe	360
Qy	1080	TAAGCAGGCTTTTCACTTTCTCGCAACTTCAAGGCTTTCTGTGTAAACAATATCTGAA	1139
Db	360	rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysGlnGlnTyrLeuHi	380
Qy	1140	CCTTTACCCCGTTGCCCGCAACGGTCCGCTCTCTGCAAGTGTCTGACCGAACCC	1199
Db	380	sLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr	400
Qy	1200	CACGTGATGGGCTTGGCCATAGCCATAGCCATAGCCATAGCCATAGCCATAGCCATAG	1259
Db	400	oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe	420
Qy	1260	GCCGATCCATCTGCGGAACCTCTAGCAGCTTGTGTTGCTCGCAGCGGCTTGGAGCAA	1319
Db	420	uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy	440
Qy	1320	ACTTATCGGAACCGACAACTCTGTGCTCTCTCGGAAATACACCTCTTCCATGGCT	1379
Db	440	sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe	460
Qy	1380	GCTAGGTGTGCTGCAACTGATGATCTGCGGGGACGCTCTTGTCTAGCTCCGCTGGC	1439
Db	460	uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl	480
Qy	1440	GCTGAATCCCGCGACGACCGCTCTCGGGCGGCTTGGGGCTCTACCGTCCCTTCTTCA	1499
Db	480	aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi	500
Qy	1500	TCTGCCCTTCCGGCCGACACCGGCGGCTCTCTTTACGCGGTCTCCCGTATGTC	1559
Db	500	sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr	520
Qy	1560	TTCATCTCTCGCGACCGTGTGCACTTGGTTCACCTCTGCACTCTGCACTGCGATGAGCACC	1619
Db	520	oSerHisLeuProSerArgValHisPheAlaSerProLeuHisValAlaIleTrpArgProPr	540
Qy	1620	G 1620	
Db	540	o 540	

RESULT 10

Q5KR39_HPBVO PRELIMINARY; PRT; 843 AA.

ID Q5KR39; HPBVO PRELIMINARY; PRT; 843 AA.

AC Q5KR39; HPBVO PRELIMINARY; PRT; 843 AA.

DT 10-MAY-2005 (TrEMBLrel. 30, Created)

DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)

DE 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)

DE Polymerase.

GN Name=P;

OS Hepatitis B virus.

OC Viruses; Retro-transcribing viruses; Hepadnaviridae;

OC Orthohepadnavirus.

OC NCBI_TaxID=10407;

OX [1]

RN NUCLEOTIDE SEQUENCE.

RP Sun X., Rokuhara A., Tanaka E., Gad A., Mutou H., Matsumoto A.,

RA Yoshizawa K., Kiyosawa K.;

RT "Nucleotide Mutations Associated With Hepatitis B e Antigen

RT Negativity.";

RL J. Med. Virol. 76:170-175 (2005).

RN [2]

RN NUCLEOTIDE SEQUENCE.

RP Rokuhara A.;

RA Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.

RL EMBL; AB198077; BAD86600.1; -; Genomic_DNA.

DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0004519; F:endonuclease activity; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0004523; F:ribonuclease H activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006278; F:RNA-dependent DNA replication; IEA.
 DR InterPro; IPR001462; DNAPol_viral_C.
 DR InterPro; IPR000201; DNAPol_viral_N.
 DR InterPro; IPR000477; RVtse.
 DR Pfam; PF00336; DNA_pol_viral_C; 1.
 DR Pfam; PF00242; DNA_pol_viral_N; 1.
 DR Pfam; PF00078; RVt_1; 1.
 DR ProDom; PD000814; DNAPol_viral_C; 1.
 SQ SEQUENCE 843 AA; 94446 MW; 80CD75F48E81AC2F CRC64;

Alignment Scores:

Pred. No.:	1.4e-200	Length:	843
Score:	2806.00	Matches:	526
Percent Similarity:	98.15%	Conservative:	5
Best Local Similarity:	97.23%	Mismatches:	9
Query Match:	47.50%	Indels:	2
DB:	2	Gaps:	0

US-10-761-006a-1 (1-3215) x Q5KR39_HPBVO (1-843)

QY 1 CTCACAAATTCACCAAGCTCTGCTAGATCCAGGGTGAGGGGCTATATTTCTCTGC 60
 |||||
 DB 304 LeuHisAsnIleProProSerSerAlaAtrpGlnSerGluGlyProIleLeuSerCys 323
 |||||

QY 61 TGTGGCTCCAGTCCGGAACAGTAAACCTGTTCGACTACTGCTCTCCCATATCGTC 120
 |||||
 DB 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
 |||||

QY 121 AATCTTCTGAGACTGGGACCTGACCGACATCGAGAACACATCAGATTCTCT 180
 |||||
 DB 344 AsnLeuLeuGlnAspTrpGlyProCysThrGluHisGlyHisAsnIleArgIlePro 363
 |||||

QY 181 AGGACCCCTCTCGTGTATACAGCGGGGTTTTCTGTTGACAAAGATCCTCAATACC 240
 |||||
 DB 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
 |||||

QY 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCAATTTCTAGGGGAGACCCACGTGTC 299
 |||||
 DB 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
 |||||

QY 300 CTGGCCAAATTCGAGTCCCGAACCTCCATCACTCAACCACTCTGTCTCCCAATT 359
 |||||
 DB 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 423
 |||||

QY 360 GTCCTGGCTATCCTGATGTGCTGGGGGTTTTATCATATTCCTCTCATCTCTGTCG 419
 |||||
 DB 423 uSerTrpLeuSerLeuAspValSerAlaPheTyrHisIleProLeuHisProAlaAl 443
 |||||

QY 420 TATGCTCATCTTCTTGTGGTTCTTCTGAGCTACCAAGTATGTTGCCGGTTGCTTC 479
 |||||
 DB 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
 |||||

QY 480 TACTTCCAGGAACATCAACACAGACCGGGGCATCGACCTGACGACTCTCTGCTC 539
 |||||
 DB 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
 |||||

QY 540 AAGGAACCTACGTTTCCCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
 |||||
 DB 483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
 |||||

QY 600 GTATTCCTCCATCATCTCGGCTTTTCGAAAGTTCCTATCGGAGTGGGCTCAGTCC 659
 |||||
 DB 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
 |||||

QY 660 GTTTCCTCGGCTCAGTTTACTAGTCCATTGTTTTCAGTGGTTCGTAGGGCTTCCCCCA 719
 |||||

RESULT 11

Q7TDR3 HPBVO

ID Q7TDR3 HPBVO PRELIMINARY; PRT; 843 AA.

AC Q7TDR3;

DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DB 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
 |||||

QY 720 CTGTTTGGCTTTTCAGTTATATGATGATGTGTTATTTGGGGCGGAAGTCTGTCAACATCT 779
 |||||

DB 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
 |||||

QY 780 TGAGTCCCTTTTACCTCTATTTACCAATTTCTTTTGTCTTTGGGTATATCATTTAAACCC 839
 |||||

DB 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
 |||||

QY 840 TAATAAAACCAACGTTGGGGTACTCCCTTAACTTCTGCGGATATGTAATTCGAAGTTG 899
 |||||

DB 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
 |||||

QY 900 GGGTACTTTTACCGCAGGAACATATTTCTATAAATCAAGCAATGTTTTCGAAAACTGCC 959
 |||||

DB 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
 |||||

QY 960 TGTAAATAGACCTATTGATTGGAAGTATGTCAAGAATTGTGGTCTTTTGGGCTTTGC 1019
 |||||

DB 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
 |||||

QY 1020 TGCCTCTTTTACACAATGTGGCTATCTCTGCTTGATGCTTTTATATGATATATCAATC 1079
 |||||

DB 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
 |||||

QY 1080 TAAGCAGGCTTTTCACTTTCTCGCCAACTTCAAGGCTTTTCTGTGTAAAACAATATCTGAA 1139
 |||||

DB 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 683
 |||||

QY 1140 CCTTTACCCGTTTCCCGGCAAGGTCCTGCCAAGTGTGTTTGTGACGCAACCC 1199
 |||||

DB 683 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
 |||||

QY 1200 CACTGATCGGGCTTCGCGCATAGGCCATCAGCGCATGGCTGGAACTTTCTGGCTCTCT 1259
 |||||

DB 703 oThrGlyTrpGlyLeuAlaValGlyHisArgMetArgGlyThrPheValAlaProLe 723
 |||||

QY 1260 GCGATCCCATCTCGGAACTCCTAGCAGCTGTTTGTCTCGCAGCGGCTCTCGAGCAA 1319
 |||||

DB 723 uProIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLys 743
 |||||

QY 1320 ACTTATCGGAACCGCAACTCTGTTCTCTCTCGGAATAATACACCTCTTTCATGGCT 1379
 |||||

DB 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
 |||||

QY 1380 GCTTAGGCTGTGTCGCAACTGGATCCTCGCGGAGCTCTTGTCTAGTCCCGTCGGC 1439
 |||||

DB 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
 |||||

QY 1440 GCTGAATCCCGGAGACGCCGCTCTCGGGGCCGTTTGGGGCTCTACGCTCCCTTCTTCA 1499
 |||||

DB 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
 |||||

QY 1500 TCTGCGTTTCCGGCCGACCAACCGGGCGCACTCTCTTTTACGCGGTCTCCCGGTATGTC 1559
 |||||

DB 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
 |||||

QY 1560 TTCTCATCTGCGGACCGGTGTCACCTTCCTTCACCTCTGCGCTGCGCATGGAGACACC 1619
 |||||

DB 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
 |||||

QY 1620 G 1620
 |||||

DB 843 O 843

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Polymerase protein.
 OS Hepatitis B virus.
 OC Viruses; Retroviral viruses; Hepadnaviridae; Orthohepadnavirus.
 ON NCBI_TaxID=10407;
 RX [1]
 RP NUCLEOTIDE SEQUENCE
 RA Liu C.-J., Chen P.-J., Lai M.-Y., Kao J.-H., Chen D.-S.;
 RL Submitted (Oct-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY167096; AA041314.1; -; Genomic DNA.
 DR GO: 0003677; F: DNA binding; IEA.
 DR GO: 0003887; F: DNA-directed DNA polymerase activity; IEA.
 DR GO: 0004519; F: endonuclease activity; IEA.
 DR GO: 0016787; F: hydrolase activity; IEA.
 DR GO: 0004523; F: ribonuclease H activity; IEA.
 DR GO: 0003723; F: RNA binding; IEA.
 DR GO: 0003964; F: RNA-directed DNA polymerase activity; IEA.
 DR GO: 0016740; F: transferase activity; IEA.
 DR GO: 0006278; P: RNA-dependent DNA replication; IEA.
 DR InterPro; IPR001462; DNaPol_viral_C.
 DR InterPro; IPR00201; DNaPol_viral_N.
 DR InterPro; IPR00477; RVTse.
 DR Pfam; PF00336; DNA_pol_viral_C; 1.
 DR Pfam; PF00242; DNA_pol_viral_N; 1.
 DR Pfam; PF00078; RVT_1; 2.
 DR ProDom; PD000814; DNaPol_viral_C; 1.
 SQ SEQUENCE 843 AA; 94350 MW; E033A0F3BD9963AB CRC64;

 Alignment Scores:
 Pred. No.: 1,98e-200 Length: 843
 Score: 2804.00 Matches: 526
 Percent Similarity: 98.15% Conservative: 5
 Best Local Similarity: 97.23% Mismatches: 9
 Query Match: 47.47% Indels: 2
 DB: 2 Gaps: 0

 US-10-761-006a-1 (1-3215) x Q7TDR3_HPBV0 (1-843)
 QY 1 CTCACAACTCCACCAAGCTCTGCTAGATCCACGGGTGAGGGGCTATATTTCTCTGC 60
 DB 304 LeuGlnAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 323
 QY 61 TGGTGGCTCCAGTCCGGAACAGTAACCTGCTCCGACTACTGCTCTCCCATATGTC 120
 DB 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
 QY 121 AATCTTCGAGGACTGGGGACCTCGACCAACATGAGAACACAAATCAGGATTCCT 180
 DB 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
 QY 181 AGGACCCCTGCTCGTGTACAGGGGGGTTTTCTCGTTGACAGAAATCCTCACATACC 240
 DB 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
 QY 241 GCAGGTCTAGACTCTGCTCAATTTCTAGGGGAGCACCACGCTGTC 299
 DB 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
 QY 300 CTGGCAAAATTCGAGTCCCAACCTCCAACTCACTACCAACCTCTTGCTCTCAATTT 359
 DB 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 423
 QY 360 GTCTGGCTATCGCTGGATGTCTGCGGGGTTTTATCATATTCCTCTTCTCTGCTGC 419
 DB 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
 QY 420 TATGCTCATCTCTCTGTTCTCTGACTACCAAGGTATGTTGCCCTTTGCTCTC 479
 DB 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaAlaArgLeuSer 463
 QY 480 TACTTCCAGAACATCAACACACGACGGGGCCATGCAAGACCTGACGACTCTCTGCTC 539
 DB 463 rThrSerArgAsnIleAsnHisGlnHisIleGlyAlaMetGlnAspLeuHisAspSerCysSe 483

QY 540 AAGGAAACTCTACGTTTCCCTCTGTTGCTGTACAAAAACCTTCGGACGGAACCTGCACTT 599
 DB 483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
 QY 600 GTATTCCTCCATCCCATCTCTGGCTTCCGACAGATTCCTATGGAGTGGGCTCAGTCC 659
 DB 503 uTyrSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
 QY 660 GTTTCCTCTGGCTCAGTTACTAGTCCATCTCTCAGTGGTTCGTAGGCTTCCGCCA 719
 DB 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProH 543
 QY 720 CTGTTGGCTTTCAGTTATATGATGATGATGGGTATTTGGGGCGGAAGTCTGTACAACATCT 779
 DB 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
 QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTCTTTGGGTATACATTTAAACCC 839
 DB 563 uGluSerLeuTyrThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
 QY 840 TAATAAACCAACAGTTGGGCTACTCCCTTAACCTCAAGCAATGTTTTGAAACACTGCC 899
 DB 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
 QY 900 GGGTACTTTTACCGCAGGAACATATTGACTATAAACTCAAGCAATGTTTTGAAACACTGCC 959
 DB 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
 QY 960 TGTAATAGACCTATTGATTTGGAAGATGTCATAAGAAATTTGGGTCTTTTGGGCTTTC 1019
 DB 623 oValAsn***ProIleAspTrpLysValCysGlnLysIleValGlyLeuLeuGlyPheAl 643
 QY 1020 TGCCCTTTTACACAATGTGGCTATCTCTGCTGATGCTCTTATATCATCATATACAAATC 1079
 DB 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
 QY 1080 TAAGACGGCTTTCACCTTTCTCGCAACTTACAAGGCCTTTCTGTGTAAACAATATCTGAA 1139
 DB 663 rLysGlnAlaPheThrPheSerProThrTyrLysThrPheLeuCysLysGlnTyrLeuAs 683
 QY 1140 CTTTATACCCCTGTCGGCAACCGTCCGGTCTCTGCAAGTGTGTGTGACGCAACCCC 1199
 DB 683 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
 QY 1200 CACTGATGGGCTTGGCCATAGCCCATAGCGCATGGCTGGAACTTTCTGGCTCTCT 1259
 DB 703 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 723
 QY 1260 GCCGATCCATCTCGCAACTCTTAGCAGCTTGTGTTGCTCGCAGCGGTCTGGAGCAAA 1319
 DB 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 743
 QY 1320 ACTTATCGGAACCAACTCTGTGCTCTCTCGGAATATACCTCTCTTTCATGCT 1379
 DB 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
 QY 1380 GCTAGGTGTCTGCCAAGTCTCTGCGGGAGCTCTCTTCTCTACGTCCTGCTCGCCG 1439
 DB 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
 QY 1440 GCTCAATCCCGGACGACCGCTCTCGGGGCGGTGTTGGGCTCTACCGTCCCTCTTCTCA 1499
 DB 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
 QY 1500 TCTGCCGTTCGGCGGACCAACCGGGCGCCTCTCTCTTTTACCGGCTCTCCCGTATGTGCC 1559
 DB 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
 QY 1560 TTCTCATCTGCCGACCGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1619
 DB 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaAlaTrpArgProPr 843


```

Db      763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyValProSerAl 783
Qy      1440 GCTGAATCCCGCGGACGACCGCTCTCGGGCCGCTTTCGGGCTCTACCGTCCCTTCTTCA 1499
Db      783 aLeuAsnProAlaAspAspProSerAsgGlyArgLeuGlyLeuTyArgProLeuLeuHi 803
Qy      1500 TCTGCCGTTCCGGCCGACCGGCGGCGCACTCTCTTTAGCGGTCTCCCGGTATGTGCC 1559
Db      803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyAlaValSerProSerValPr 823
Qy      1560 TTCTCATCTCGCGACCGTGTGCACTTTCGCTTTCACCTCTGCACCTCGCATCGGAGACCAACC 1619
Db      823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
Qy      1620 G 1620
Db      843 o 843

RESULT 13
Q4PD65_HPBV0 PRELIMINARY; PRT; 843 AA.
AC Q4PD65;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Polymerase.
GN Name=P;
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=HK477;
RX PubMed=15897987; DOI=10.1086/430324;
RA Chan H.L., Tsui S.K., Tse C.H., Ng E.Y., Au T.C., Yuen L.,
RA Bartholomeusz A., Leung K.S., Lee K.H., Locarnini S., Sung J.J.; of
RT "Epidemiological and virological characteristics of 2 subgroups of
RT hepatitis B virus genotype C.";
RL J. Infect. Dis. 191:2022-2032(2005).
DR EMBL: D0089798; AAZ05298.1; -; Genomic DNA.
SQ SEQUENCE 843 AA; 94318 MW; 0C749903F281504A CRC64;

Alignment Scores:
Pred. No.: 2,35e-200 Length: 843
Score: 2803.00 Matches: 525
Percent Similarity: 98.34% Conservative: 7
Best Local Similarity: 97.04% Mismatches: 8
Query Match: 47.45% Indels: 2
DB: 2 Gaps: 0

US-10-761-006a-1 (1-3215) x Q4PD65_HPBV0 (1-843)
Qy 1 CTCGCAACATTCACCAAGCTCTCTAGATCCCGAGGTGAGGGCTATATTTCTCTGC 60
Db 304 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 323
Qy 61 TGGTGGCTCCAGTCCGGAACAGTAAACCTGTTCCGACTACTGCTCTCCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrcysLeuThrHisIleVal 343
Qy 121 AATCTTCTCGAGGACTCGGGACCCCTGCACCGAACATGAGAACACAAATCAGGATTCCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
Qy 181 AGACCCCTGCTCGTGTGTACAGCGCGGTTTTCTCGTTGACAAATCCTCACAAATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
Qy 241 GCAGAGTCTAGACTCTG-GTGACTTCTCAATTTTCTAGGGGGAGCACCCACCGTGTTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403

```

```

Qy      300 CTGGCCAAATTCGAGTCCCAACCTCAATCACTCACCAACCTCTGTCTCTCCAAATTT 359
Db      403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
Qy      360 GTCTCTCGCTATCCCTGGATGTGTCTCGGGGTTTTATCATATTCTCTTTCATCTCTGCTGC 419
Db      423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyHisIleProLeuHisProAlaAl 443
Qy      420 TATGCTCATCTCTTCTGTTGGTCTCTCGACTACCAAGGTATGTTGCCCGTTTGTCTCTC 479
Db      443 aMetProHisLeuLeuValGlySerSerGlyLeuProGlyTyValAlaArgLeuSerSe 463
Qy      480 TACTTCCAGGAACATCAACCAACGACGCGGGCCATGCAAGACCTGACGACTCTCTCTCTC 539
Db      463 rThrSerArgAsnIleAsnTyrglnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
Qy      540 AAGGAAACTCTACGTTTCCCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
Db      483 rArgAsnLeuTyValSerLeuLeuLeuTyLysThrPheGlyArgLysLeuHisLe 503
Qy      600 GTATTCCCATCCCATCATCTCTGGGCTTTCGCAAGATTCTATGGAGTGGGCTCAGTCC 659
Db      503 uTySerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Qy      660 GTTCTCTCTCGCTCAGTTTACTAGTGCATTGTTCAGTGTGTCTGTAGGGGCTTTCCCCCA 719
Db      523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 543
Qy      720 CTGTTTGGCTTTTTCAGTTATATGATGATGATGATGATGATGATGATGATGATGATGAT 779
Db      543 sCysLeuAlaPheSerTyMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
Qy      780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 839
Db      563 uGluSerLeuTyThrAlaIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
Qy      840 TATATAAACCAACGTTGGGCTACTCCCTTAACCTTCACTGGATATGTAATTTGGAAGTTG 899
Db      583 oAsnLysThrLysArgTrpGlyTySerLeuAsnPheMetGlyTyValIleGlySerTr 603
Qy      900 GGGTACTTTTACCGCAGAACATATGTTACTAAACTCAAGCAATGTTTTTCAAAAACCTGCC 959
Db      603 pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr 623
Qy      960 TGTAAATAGACCTTATTCATTGGAAGTATCTCAAGATTTGCGGTCTTTTGGGCTTTTGC 1019
Db      623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
Qy      1020 TGCCCTTTTACACAATGTGCTATCTGCTTCTGATGCTTATATGATGATGATGATGATGATG 1079
Db      643 aAlaProPheThrGlnCysGlyTyProAlaLeuMetProLeuTyAlaCysIleGlnSe 663
Qy      1080 TAAGCAGGCTTTTCACTTTTCTCGCAACTTACAAGGCTTTTCTGTGTAAACATATCTGAA 1139
Db      663 rLysGlnAlaPheThrPheSerProThrTyLysAlaPheLeuLeuCysLysGlnTyLeuAs 683
Qy      1140 CTTTATACCTTCCCGGCAACCGTCCGCTCTCTGCAAGTGTGTTGCTGACGCAACCC 1199
Db      683 nLeuTyProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
Qy      1200 CACTGGATGGGCTTGGCCATAGGCCATCAGCGCATGCTGGAACCTTTCTTGCGTCTCTCT 1259
Db      703 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 723
Qy      1260 GCCGATCCATCTCGGAACTCTTAGCAGCTTGTGTTTGTCTCGCAGCCGGTCTGAGAGAAA 1319
Db      723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 743
Qy      1320 ACTTATCGGAACCGCAACTCTGTGCTCTCTCGGAATACACCTCTCTTCCATGGCT 1379
Db      743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyThrSerPheProTrpLe 763

```

```
QY 1380 GCTAGGGTGTGCTGCCAAGTGGATCTCTGCGCGGAGCGTCTTTGTCTAGCTCCCGTCGGC 1439
DB 763 uLeuGlyCyAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
QY 1440 GCTGAATCCCGCGGACGACCGCTCTGCGGCGCGTTTGGGGCTCTACCGTCCCGTCTTCA 1499
DB 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuPheArgProLeuLeuHl 803
QY 1500 TCTGCCGTTCCGCGCAGCACCGGGCGCACCTCTCTTTACCGGCTCTCCCGTATGTCC 1559
DB 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
QY 1560 TTCTCATCTGCCGACCGTGTGCACCTTCTGCTTCTGACCTCTGCACGTGGCAGACCA 1619
DB 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
QY 1620 G 1620
DB 843 o 843
```

RESULT 14

```
ID Q81165_HPBVO PRELIMINARY; PRT; 838 AA.
AC Q81165;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE DNA polymerase.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=subtype adr;
RX MEDLINE=93096607; PubMed=1461746;
RA Mukaid M.;
RT "The complete nucleotide sequence of hepatitis B virus, subtype adr
RT (SRAR) and phylogenetic analysis.";
RL Nucleic Acids Res. 20:6105-6105(1992).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=subtype adr;
RA Uchida T., Aye T.T., Shimojima M., Gotoh K., Shikata T.;
RT "Full-length nucleotide sequence of a hepatitis B virus (HBV) mutant
RT isolated from a patient with acute hepatitis who did not exhibit
RT serological markers for HBV infection.";
RL Int. Hepatol. Commun. 2:70-73(1994).
DR EMBL; D16665; BAA04072.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001462; DNAPol_viral_C.
DR InterPro; IPR000201; DNAPol_viral_N.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00078; RVT_1; 2.
DR ProDom; PD000814; DNAPol_viral_C; 1.
SQ SEQUENCE 838 AA; 93902 MW; A686F42940C3B806 CRC64;
```

Alignment Scores:

Pred. No.:	2,78e-200	Length:	838
Score:	2802.00	Matches:	525
Percent Similarity:	98.33%	Conservative:	6
Best Local Similarity:	97.22%	Mismatches:	8
Query Match:	47.44%	Indels:	2
DB:	2	Gaps:	0

```
US-10-761-006A-1 (1-3215) x 081165_HPBVO (1-838)
QY 4 CACAACATTCCACCAAGCTCTCTAGATCCAGGGTGAGGGCGCTATATTTCTCTGCTGG 63
DB 300 HisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCysTrp 319
QY 64 TGGCTCCAGTTCCGGAACAGTAAACCTGTTCCGACTACTGCTCTCCCATATCGTCAAT 123
DB 320 TrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAsn 339
QY 124 CTTCTCGAGGACTGGGGACCTCGCACCAACATGGAGAGACCAACATCAGGATTCCTAGG 163
DB 340 LeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArg 359
QY 184 ACCCTGCTCGTGTGTACAGCGGGGGTTTTCTCGTTGACAAGAATCCTCACATACCGCA 243
DB 360 ThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThrThr 379
QY 244 GAGTCTAGACTCTG- GTGACTTCTCTCAATTTTCTAGGGGAGAGACCCACGTTGCTCTG 302
DB 380 GluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-SerTr 399
QY 303 GCCAAATTCGCAGTCCCAACCTCCAATCACTCAACCACTCTCTGCTCCCAATTTGTC 362
DB 399 pProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLeuSe 419
QY 363 CTGGCTATCGCTGGATGTCTCGCGCGTTTTATCATATTCCTTTCATCCTGCTGCTAT 422
DB 419 rTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAla 439
QY 423 GCCTCATCTTCTGTGTCTCTGACTACCAAGTATGTTGCCCGTTGCTCTCTAC 482
DB 439 tProHisLeuLeuValGlySer-SerGlyLeuProArgTyrValAlaArgLeuSerSerTh 459
QY 483 TTCCAGGAACATCAACCCAGCAGCGGGCCATGCAAGACCTGCACGACTCTCTGCTCAAG 542
DB 459 rSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSerAr 479
QY 543 GAAACTCTACGTTTCCCTCTTGTGTGTACAAAACCTTCGGACGGAAACTGCACATTGA 602
DB 479 gAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeuTy 499
QY 603 TTCCCATCCATCATCTCGGCTTTTCGGAAGATTCCTATGGAGTGGGGCTCAGTCCGTT 662
DB 499 rSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPh 519
QY 663 TCTCTCGGCTCAGTTTACTAGTGCCATTTGTTTCAGTGGTTCTGTAGGGCTTTCCCCCACTG 722
DB 519 eLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHisCy 539
QY 723 TTTGGCTTTTCAGTTATATGATGTGTGTTATTTGGGGGCGAAGTCTGTACAACATCTTGA 782
DB 539 sLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGl 559
QY 783 GTCCCTTTTACCTCTATTACCAATTTCTTTTCTTTTGGGTATACATTTAACCTTAA 842
DB 559 uSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAs 579
QY 843 TAAACCAAAACGTGGGGCTACTCCCTTAACCTTCATGGGATATGTAATTTGGAAGTTGGGG 902
DB 579 nLysThrLysArgTrpGlyTyr-SerLeuAsnPheMetGlyTyrValIleGlySerTrpGl 599
QY 903 TACTTTACCGCAGGAACATATTGTACTAAACCTCAAGCAATGTTTTCGAAAACTGCTGCT 962
DB 599 yThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuProVa 619
QY 963 AAATAGACCTATTGATTGGAAGATATGTCAAAGATTGGGCTCTTTGGGCTTTCCTGTC 1022
DB 619 lAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAl 639
QY 1023 CCCTTTTACCAATGTGGCTATCTCGCTTGTATGCTCTTATATGATGATATACAACTTAA 1082
```


Db 639 aProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerLy 659
Qy 1083 GCAGGCTTTACATTTCTCGCAACTTACAGGCTTTCTGTGTAAACATATCTGAACCT 1142
Db 659 sGlnAlaPheThrPheSerProThrTyrLyAlaPheLeuCysGlnGlnTyrLeuHisLe 679
Qy 1143 TTACCCCGTTGCGCGGCAACGGTCCGGTCTCTGCAAGTGTCTGCTGACCAACCCCCAC 1202
Db 679 uTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProTh 699
Qy 1203 TGGATGGGCTTGCCATAGCCATCAGCCATGCTGGACCTTTCTGCTCTCTGCTGCC 1262
Db 699 xGlyTyrGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLeuPr 719
Qy 1263 GATCCATCTGCGGAATCTCTAGAGCTCTTTTGTCTGCGAGCGGTCTGAGCAAACT 1322
Db 719 oIleHisThrAlaGlnLeuAlaAaCysPheAlaArgSerArgSerGlyAlaHisLe 739
Qy 1323 TATCGGAACCGAACACTCTGTGTCTCTCTCGGAATACACCTCTTCCATGGCTGCT 1382
Db 739 uIleGlyThrAspAsnSerValValLeuSerArgLyThrSerPheProTyrPheLeu 759
Qy 1383 AGGTGTGCTGCCAACTGGATCTTCGCGGAGCTCTTTGTCTACCTCCGCTCGGCGCT 1442
Db 759 uGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAlaLe 779
Qy 1443 GAATCCCGGAGACACCGTCTCGGGCGGCTTTGGGGCTCTACGCTCCCTCTTCTCATCT 1502
Db 779 uAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHisLe 799
Qy 1503 GCCGTTCGCGGACCGACCGGGCGACCTCTCTTTACGCGGTCTCCCGCTATGCGCTTC 1562
Db 799 uProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValProSe 819
Qy 1563 TCATCTGCGGACCGCTGTCACCTCTCGCTTCACTCTCGACGTGCGATGGAGACCCG 1620
Db 819 xHisLeuProSerArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 838

RESULT 15
O42041 HPBV0
ID O42041 HPBV0 PRELIMINARY; PRT; 843 AA.
AC O42041;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE DNA polymerase.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RA Sato C.;
RT "Sequential changes in full-length genomes of hepatitis B virus
RT accompanying acute exacerbation of chronic hepatitis B.";
RL J. Hepatol. 25:787-794 (1996).
DR EMBL; D50518; BAA23442.1; -; Genomic DNA.
DR EMBL; D50517; BAA23435.1; -; Genomic DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006278; F:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001462; DNAPol_viral_C.
DR InterPro; IPR000201; DNAPol_viral_N.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.

DR Pfam; PF00078; RVT_1; 2.
DR ProDom; PD000814; DNAPol_viral_C; 1.
SQ SEQUENCE 843 AA; 94462 MW; 1BCCF79639BB140D CRC64;
Alignment Scores:
Pred. No.: 2,79e-200 Length: 843
Score: 2802.00 Matches: 525
Percent Similarity: 98.52% Conservative: 8
Best Local Similarity: 97.04% Mismatches: 7
Query Match: 47.44% Indels: 2
DB: 2 Gaps: 0
US-10-761-006a-1 (1-3215) x O42041_HPBV0 (1-843)
Qy 1 CTCACAACTTCCACCAAGCTCTGTAGATCCAGGGTGAGGGGCTATATTTTCTCTGC 60
Db 304 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGlnGlyProIlePheSerCys 323
Qy 61 TGTGTGCTCCAGTTCGGAACAGTAAACCTGTGTCGACTACTGCTCTCCCATATGCTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLyProCysSerAspTyrCysLeuThrHisLeVal 343
Qy 121 AATCTTCTCGAGACTGGGGACCTCGACCGAATCGGAGAACACACATCATGAGATTCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGlnHisAsnIleArgIlePro 363
Qy 181 AGGACCCCTGCTGTTTACAGCGCGGTTTTCTCGTTCACAAGATCTCTCAATATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLyAsnProHisAsnThr 383
Qy 241 GCAGACTGTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGGAGCACCCACGTGTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
Qy 300 CTGGCCAAATTCGAGTCCCACTCCGATCACTCACCAACCTCTGTCTCTCAATTT 359
Db 403 rTrpProLyPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
Qy 360 GTCCTGGCTATCGCTGGATGTGTGCGGGTTTTATCATATTCCTCTCTCATCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
Qy 420 TATGCTCTCATCTTCTTGTGTTCTCTGAGCTACCAAGGTATGTTGCCGTTGTCCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
Qy 480 TACTTCCAGGAACATCAACACAGCAGCGGGCCATCGAAGCTCGACGACTCTCTGCTC 539
Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisGluSerCysSe 483
Qy 540 AAGGAACTCTAGCTTTCTCTTGTGCTGTACAAAACCTTCGAGCGGAACTGCACCTT 599
Db 483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLyLeuHisLe 503
Qy 600 GTATTCCTCCATCCATCATCTGGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
Db 503 uTyrSerHisProIleIleLeuGlyPheArgLyAlaProMetGlyValGlyLeuSerPr 523
Qy 660 GTTTCCTCTGGCTCAGTTTACTAGTGCATTTGTTTCACTGTTGTTGTTAGGCTTTCCCCA 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 543
Qy 720 CTGTTTGGCTTTCAGTTATATGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 779
Db 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysserValGlnHisLe 563
Qy 780 TGAGTCCCTTTTACCTCTATTACCAATTTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTG 839
Db 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeAsnPr 583
Qy 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACCTTCACTGATGATGATGATGATGATGATG 899
Db 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603

QY 900 GGCTACTTTACCGCAGACATATTGCTACATACTAAGCAATGTTTCGAAACTGCC 959
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
 QY 960 TGTAAATAGACCTATTGATTGGAAGTATGTCAAAGAAATTGTGGGTCTTTTGGGCTTTGC 1019
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
 QY 1020 TGCCCTTTTACACAATGTGGCTATCTGCTTGTATGCTCTTTATATGCAATGATACAATC 1079
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
 QY 1080 TAGCAGGCTTTCACCTTCTCGGCAACTTACAGGCTTCTCTGTGTAAACAATATCTGAA 1139
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysGlnGlnTyrLeuHi 683
 QY 1140 CCTTTACCCCGTTGCCGCGCAACGGTCCGGTCTCTGCAAGTGTGCTGACGCAACCCC 1199
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 683 sLeuTyrProValAlaArgGlnArgSerGlyValCysGlnValPheAlaAspAlaThrPr 703
 QY 1200 CACTGATGGGCTTGGCCATAGCCCATCAGCCCATGGCTGGAACCTTTCTGGCTCTCT 1259
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 703 oThrGlyTrpGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe 723
 QY 1260 GCCGATCCATCTGCGGAACCTCTAGCAGCTTGTGTTGCTCGCAGCGGCTCGGACAAA 1319
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 723 uProIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLy 743
 QY 1320 ACTTATCGGAACCGACAACCTCTGTGCTCTCTCGGAAATACACCTCTTTCATGGCT 1379
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProIrpLe 763
 QY 1380 GCTAGGCTGTGCTGCCAAGTGGATCTGCGCGGACGCTCTTTGTCTACGTCCCGTCGCG 1439
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
 QY 1440 GCTGAATCCCGCGGACGACCGCTCTCGGGCCGTTTGGGGCTCTACCGTCCCGCTTCTCA 1499
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
 QY 1500 TCTGCCGTTCCGGCCGACACGCGGCGCACCTCTCTTTACGGGTCTCCCGGTATGTGCC 1559
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
 QY 1560 TTCTCATCTGCGGACGCTGTGCACTTCGCTTCACTCTGCACTGCGATGGAGACCACC 1619
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
 QY 1620 G 1620
 Db |
 843 o 843

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: December 1, 2005, 23:16:00 ; Search time 94.5 Seconds
(without alignments)
5625.446 Million cell updates/sec

Title: US-10-761-006A-1
Perfect score: 5907
Sequence: 1 CTCACACATTCACCAAG.....CCTCAGGCCACGAGTGGA 3215

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0
Searched: 572060 seqs, 82675679 residues
Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/p/US10761006/runat_01122005_113953_26895/app_query.fasta_1.3399
-DB=Issued Patents AA -OFMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10761006_@CGN_1_140 @runat_01122005_113953_26895 -NCFU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2927	49.6	843	2	US-09-719-528A-2
2	2927	49.6	843	2	US-10-209-264-2
3	2807	47.5	843	2	US-08-591-502B-45
4	2807	47.5	843	2	US-08-591-502B-59
5	2188	47.2	843	2	US-08-591-502B-47
6	2773	46.9	845	2	US-08-591-502B-46
7	2759	46.7	843	2	US-08-591-502B-48
8	2759	46.7	845	1	US-08-416-950-11
9	2759	46.7	845	1	US-08-469-830-11
10	2741	46.4	845	2	US-08-591-502B-11
11	2734.5	46.3	842	2	US-08-591-502B-51
12	2729.5	46.2	842	2	US-08-591-502B-50

13	2652	44.9	832	2	US-08-591-502B-61	Sequence 61, Appl
14	2646	44.8	832	2	US-08-591-502B-62	Sequence 62, Appl
15	2645	44.8	843	2	US-08-591-502B-56	Sequence 56, Appl
16	2643	44.7	832	2	US-08-591-502B-60	Sequence 60, Appl
17	2643	44.7	843	2	US-08-591-502B-55	Sequence 55, Appl
18	2637	44.6	832	2	US-08-591-502B-64	Sequence 64, Appl
19	2633	44.6	843	2	US-08-591-502B-53	Sequence 53, Appl
20	2615	44.3	843	2	US-08-591-502B-54	Sequence 54, Appl
21	2608	44.2	845	2	US-08-591-502B-57	Sequence 57, Appl
22	2607	44.1	832	2	US-08-591-502B-63	Sequence 63, Appl
23	2607	44.1	845	2	US-08-591-502B-58	Sequence 58, Appl
24	2607	44.1	845	6	5196194-17	Patent No. 5196194
25	2179	36.9	730	2	US-08-591-502B-49	Sequence 49, Appl
26	1496	25.3	400	2	US-09-719-528A-3	Sequence 3, Appl
27	1496	25.3	400	2	US-10-209-264-3	Sequence 3, Appl
28	1427	24.2	277	6	5164485-2	Patent No. 5164485
29	1423	24.1	281	2	US-09-247-890-10	Sequence 10, Appl
30	1423	24.1	281	2	US-09-724-969-10	Sequence 10, Appl
31	1423	24.1	281	2	US-09-724-852-10	Sequence 10, Appl
32	1365.5	23.1	481	2	US-08-591-502B-52	Sequence 52, Appl
33	1359	23.0	281	1	US-08-105-483-214	Sequence 214, App
34	1359	23.0	281	1	US-08-709-209-214	Sequence 214, App
35	1359	23.0	281	1	US-08-458-101-214	Sequence 214, App
36	1359	23.0	389	1	US-08-105-483-216	Sequence 216, App
37	1359	23.0	389	1	US-08-105-483-219	Sequence 219, App
38	1359	23.0	389	1	US-08-709-209-216	Sequence 216, App
39	1359	23.0	389	1	US-08-709-209-219	Sequence 219, App
40	1359	23.0	389	1	US-08-458-101-216	Sequence 216, App
41	1359	23.0	389	1	US-08-458-101-219	Sequence 219, App
42	1354	22.9	281	2	US-09-247-890-12	Sequence 12, Appl
43	1354	22.9	281	2	US-09-724-969-12	Sequence 12, Appl
44	1354	22.9	281	2	US-09-724-852-12	Sequence 12, Appl
45	1354	22.9	281	2	US-09-721-480-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-719-528A-2
; Sequence 2, Application US/09719528A
; Patent No. 6558675
; GENERAL INFORMATION:
; APPLICANT: Oon, Chong Jin
; Lim, Gek Keow
; Zhao, Yi
; Chen, Wei Ning

TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
USSES THEREOF

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ladas & Parry

STREET: 26 West 61 Street

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10023

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/719,528A

FILING DATE: 30-Apr-2001

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/SG98/00046

FILING DATE: 19-JAN-1998

ATTORNEY/AGENT INFORMATION:

NAME: Mass, Clifford J.

REGISTRATION NUMBER: 30,086

REFERENCE/DOCKET NUMBER: U-013109-7

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 708-1800

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 843 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-719-528A-2

Alignment Scores:

Pred. No.: 5,09e-259 Length: 843

Score: 2927.00 Matches: 540

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 49.55% Indels: 0

DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x US-09-719-528A-2 (1-843)

QY	1	CTCCAGAACATTCACCAAGCTCTGCTAGATCCAGGGTGGGGCTATATTTCTCTGC	60
Db	304	LeuHisAsnIleProProSerSerAlaAArgSerGlnGlyGluGlyProIlePheSerCys	323
QY	61	TGTGGCTCCAGTTCCGGACAGTAAACCTGTTCCGACTACTGCTCTCCCATATCGTC	120
Db	324	TrpIrpLeuGlnPheArgAenSerIysProCysSerAspTyrCysLeuSerHisIleVal	343
QY	121	AATCTTCTCAGGACTGGGACCGCTGCACCGAATCGGAGAACACAACTCAGGATTCCT	180
Db	344	AsnLeuLeuGlnAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro	363
QY	181	AGGACCTCTCGTGTTACAGCGGGGTTTTCTCGTTGACAAAGAACTCTCAATACC	240
Db	364	ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr	383
QY	241	GCAGGCTAGACTGTGGTACTTCTCAATTTCTAGGGGAGACCCAGCTGTTCC	300
Db	384	AlaGluSerArgLeuIrpTrpThrSerLeuAsnPheLeuGlyAlaProThrCysSer	403
QY	301	TGGCCAAAATTCAGTCCCACTCCCAATCACTCACCACCTCTTGCTCTCCAATTG	360
Db	404	TrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeu	423
QY	361	TCTGGCTATCGTGATGTCTGGCGGCTTTTATCATATTCCTCTTCATCCTGCTGCT	420
Db	424	SerTrpLeuSerLeuAspValSerAlaAlaPheThrHisIleProLeuHisProAlaAla	443
QY	421	ATGCCTCATCTTCTGTTGGTCTTCTGGACTACCAAGGTATGTTGCCGTTTGTCTCT	480
Db	444	MetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSer	463
QY	481	ACTTCCAGGAACATCAACACACGAGCGGGCCATGCAAGACCTCGACGACTCTCTCTCA	540
Db	464	ThrSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSer	483
QY	541	AGAAACTCTACGTTTCCCTCTGTTGCTGTACAAAACCTTCGACGGAACCTGCACTTG	600
Db	484	ArgLysLeuTyrValSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	503
QY	601	TATTCCTCCATCATCTCTGGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCCG	660
Db	504	TyrSerHisProIleLeuGlyPheArgGlyLeuProMetGlyValGlyLeuSerPro	523
QY	661	TTTCTCTGCTCAGTTTACTAGTGCCATTTGTTGCTAGTGGTCTGTTGGGCTTTCCCCAC	720
Db	524	PheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHis	543
QY	721	TGTTTGGCTTTCAGTTATATGGATATGTTGTTATCGGGCGAAGTCTGTACAACTTT	780
Db	544	CysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeu	563
QY	781	GAGTCCCTTTTACCTTATTACCAATTTTCTTTTGTCTTTGGGTATACATTTAAACCT	840

Db	564	GluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPro	583
QY	841	AATAAAACCAAAAGTTGGGGCTACTCCCTTAACTTCATGGGATATGTAATTTGGAAGTTGG	900
Db	584	AsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrp	603
QY	901	GGTACTTTACCGGAGGAACATATTGTACTTAAACCTCAAGCAATGTTTTCGAAACATGCCT	960
Db	604	GlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPro	623
QY	961	GTAATATAGACTATTGATTGGAAGTATGTCATAAGAAATGTCGGTCTTTTGGCTTTGCT	1020
Db	624	ValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAla	643
QY	1021	GCCCTTTTACACAATGTGGCTATCCTGCTTGTATGCTCTTATATGTCATGTATACAACT	1080
Db	644	AlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSer	663
QY	1081	ARGAGGCTTTCACITTTCTGCGCAACTTACAAGGCTTCTGTGTAAACAATATCTGAAC	1140
Db	664	LysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsn	683
QY	1141	CTTTACCCCGTTGCCCGCAACGGTCCCGTCTCTGCCAAGTGTTCCTCAGCAACCCCC	1200
Db	684	LeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPro	703
QY	1201	ACTGATGGGGTGTGCCCATAGCCATCAGCCATGGCTGGAACTTCTTGGCTCTCTCTG	1260
Db	704	ThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetAlaGlyThrPheLeuAlaProLeu	723
QY	1261	CGATCCATCTACGCGGAACCTCTAGCAGCTTGTGTTTCTCGCAGCGCGTCTGGAGCAAAA	1320
Db	724	ProIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys	743
QY	1321	CTTATCGGAACCGACAACTCTGTGTCTCTCGGAATACACCTCTTCCATCGCTG	1380
Db	744	LeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProIrpLeu	763
QY	1381	CTAGGTGTGTGCCAACTGGATCTCTGCGGGGAGCTCTTTGTCTACCTCCCGTCGGCG	1440
Db	764	LeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAla	783
QY	1441	CTGAATCCCGCGGAGACACCCGCTCTGGGGCCGTTTGGGGCTCTACGTCCTCTTCTCAT	1500
Db	784	LeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHis	803
QY	1501	CTGCGGTTCCCGCCGACACCGCGGCGCACTCTCTTTACGCGGTCTCCCGTATGTCCT	1560
Db	804	LeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProTyrValPro	823
QY	1561	TCTCATCTCGCGGACCGTGTGCATCTTCCTTCCACCTCTGACGTCGCGATGGAGCCACCG	1620
Db	824	SerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro	843

RESULT 2

US-10-209-264-2

; Sequence 2, Application US/10209264

; Patent No. 6787142

; GENERAL INFORMATION:

; APPLICANT: Oon, Chong Jin

; Lim, Gek Keow

; Zhao, Yi

; Chen, Wei Ning

; TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND

; USES THEREOF

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Ladass & Parry

; STREET: 26 West 61 Street

; CITY: New York

; STATE: New York

; COUNTRY: USA

ZIP: 10023
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/209,264
FILING DATE: 31-Jul-2002
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/SG98/00046
FILING DATE: 19-JAN-1998
ATTORNEY/AGENT INFORMATION:
NAME: Mass, Clifford J.
REGISTRATION NUMBER: 30,086
REFERENCE/DOCKET NUMBER: U-013109-7
TELEPHONE: (212) 708-1800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 843 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-209-264-2

Alignment Scores:
Pred. No.: 5, 09e-259 Length: 843
Score: 2927.00 Matches: 540
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 49.55% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1 (1-3215) x US-10-209-264-2 (1-843)

Qy 1 CTCCACAACTTCCACCAAGCTCTGCTAGATCCAGGCTGAGGGCTATATTTCTCTGC 60
Db 304 LeuHisAsnIleProProSerSerAlaArgSerGlnGlyGluGlyProIlePheSerCys 323
Qy 61 TGGTGGCTCCAGTTCGGACAGTAACCTGTCGACTACTGCTCTCCCATATCTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal 343
Qy 121 AATCTTCTCGAGGACTGGGACCTGACCCGACCAATGAGAGAACACCAATCAGGATTCCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
Qy 181 AGGACCCCTGCTGCTGTACAGGCGGGTTCCTGTTGACAGAAATCCTCACATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
Qy 241 GCAGAGCTAGACTCTGGTGGACTCTCTCAATTTCTAGGGGAGCACCCACGGTTC 300
Db 384 AlaGluSerArgLeuTrpTrpThrSerLeuAsnPheLeuGlyGlyAlaProThrCysSer 403
Qy 301 TGGCCAAATTCGAGTCCCAACTCCCAATCACTCACCAACTCTGTCTCTCCAAATTG 360
Db 404 TrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeu 423
Qy 361 TCTGGCTATCTGATGTGTCTGGCGGTTTATCATATTCCTTCATCTCTCTGCT 420
Db 424 SerTrpLeuSerLeuAspValSerAlaPheTyrHisIleProLeuHisProAlaAla 443
Qy 421 ATGCCTCATCTCTTGTGGTCTTCTGAGTACCAAGTATGTTGCCGTTGCTCTCT 480
Db 444 MetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSer 463
Qy 481 ACTTCCAGGAACATCAACACAGACGAGCGGGCATGCAAGACTGACAGCTCTCTGCTCA 540
Db 464 ThrSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSer 483

Qy 541 AGGAAACTCTACGTTTCCCTCTTGTGCTGTACAAACCTTCGACGGAACATGCACTTG 600
Db 484 ArgGlyLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeu 503
Qy 601 TATTCCTATCCCATCATCTCTGGCTTTCGGAAGATCTCTATGGAGTGGCCCTCAGTCCG 660
Db 504 TyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPro 523
Qy 661 TTTCTCTGCTCAGTTTACTAGTGCATTTGTTAGTGGTTCGTAGGGCTTTCCCCAC 720
Db 524 PheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHis 543
Qy 721 TGTTCGCTTTCAGTTATATGATGTGATGTGGTATTTGGGGCGAAGTCTGTACAACTCTT 780
Db 544 CysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLeu 563
Qy 781 GAGTCCCTTTTACCTCTATTACCAATTTCTTTTCTTTGGGTATACATTTAAACCTT 840
Db 564 GluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPro 583
Qy 841 AATAAACCAACGTTGGGCTACTCCCTTAACCTTCAATGGCATATGTAATTGGAAGTTGG 900
Db 584 AsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrp 603
Qy 901 GGTACTTTACCGCAGGAACATATTGTACTAAACTCAAGCAATGTTTTCGAAACCTCCT 960
Db 604 GlyThrLeuProGlnGluHisIleValLeuLeuLysLeuLysGlnCysPheArgLysLeuPro 623
Qy 961 GTAAATAGACTATTGATTCGAAAGTATGCAAGAAATTTGGGTCTTTTGGGGCTTTGCT 1020
Db 624 ValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAla 643
Qy 1021 GCCCTTTTACAAATGCTGCTATCCCTGCTTATGATGCTTATATGATGATATCAATCT 1080
Db 644 AlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSer 663
Qy 1081 AAGCAGGCTTTCTACTTCTCGCCAACTTCAAGGCTTCTTGTGTAAACATATCTGAAC 1140
Db 664 LysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsn 683
Qy 1141 CTTTACCCCTGTTGCCCGCAACGCTCCGCTCTGCGCAAGTGTGTTGTCGACGCAACCC 1200
Db 684 LeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPro 703
Qy 1201 ACTGGATGGGCTTGGCCATAGGCATGCGCATGCTGCAACCTTCTGCTCTCTCTG 1260
Db 704 ThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetAlaGlyThrPheLeuAlaProLeu 723
Qy 1261 CCGATCCATCTCGCGAACTCTAGCAGCTTGTGTTGCTCGCAGCCGCTCTGGAGCAAAA 1320
Db 724 ProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 743
Qy 1321 CTTATCGGAACCCGAACTCTGTTGCTCTCTCGGAAATACACTCTCTTCCATGGCTG 1380
Db 744 LeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLeu 763
Qy 1381 CTAGGTTGCTGCTCCCACTGATCCTGCGGGGAGCTCTTGTCTAGTCCCGCTCGGCG 1440
Db 764 LeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAla 783
Qy 1441 CTGAATCCCGGAGCAGCCGCTCTCGGGCGGCTTGGGGCTCTACCGCTCCCTCTTCTCAT 1500
Db 784 LeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHis 803
Qy 1501 CTGCCGTTTCCGGCCGACACCGGGCGCACCTCTCTTTACGGGGTCTCCCGGTATGGCT 1560
Db 804 LeuProPheArgProThrThrThrGlyArgThrSerLeuTyrAlaValSerProTyrValPro 823
Qy 1561 TCTCATCTGCGGACCGGTGCTCATTCTGCTTCACTCTGACCTGCGACGTGAGACACCG 1620
Db 824 SerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 843

```

RESULT 3
US-08-591-502B-45
; Sequence 45, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus

```

Alignment Scores:		
Pred. No.:	5e-248	Length:
Score:	2807.00	Matches:
Percent Similarity:	98.33%	Conservative:
Best Local Similarity:	97.78%	Mismatches:
Query Match:	47.52%	Indels:
DB:	2	Gaps:

Qy	4	CACAACTTCCACCAAGCTCTGCTAGATGCCAGGGTGAGGGGCTATATTTTCTCTCTGG	63
Db	305	HisAsnIleProProSerSerAlaAArgSerGlnSerGluGlyProIlePheSerCysTtp	324
Qy	64	TGGCTCCAGTTCGGGAA CAGTAAACCCCTGGTTCGCACTACTGCTCTCCCATATCGTCAAT	123
Db	325	TtpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAsn	344
Qy	124	CTTCTCAGGACTGGGGACCTTGCACCGAA CATCGAGAA CACAACATCAGGATTCCTAGG	183
Db	345	LeuLeuGluAspTrogLpProCysThrGluHisGlyGluHisAsnIleAArgIleProAcrq	364

184	QY	ACCCCTGCTCGTGTTCACAGCGGGTTTTCTCGTTGACAGAAATCCTCACAAATACCGCA	243
365	Db	ThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrThr	384
244	QY	GAGCTCTAGACTCTG- GTGCACTTCTCTCAATTTTCTAGGGGGAGCACCCACGTGTTCCTG	302
385	Db	GlusErArgLeuValValAspPheSerGlnPheSerArgLysSerThrHisVal- SerTr	404
303	QY	GCCAAATTCGCAAGTCCCCAACCTCCAATCACATCCAACTCTTGTCTCCAAATTTGTC	362
404	Db	pProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSe	424
363	QY	CTGGCTATCGCTGGATGTGTCTGGCGGTTTTATCATATTCTCTTCATCCTGCTGCTAT	422
424	Db	rTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAlaMe	444
423	QY	GCCTCATCTTCTTTGTTGTTCTCTGAGCTACCAAGGTATGTTGCCGTTTGTCTCTAC	482
444	Db	tProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerTh	464
483	QY	TTCCAGGAACATCAACCCACGACGCGGGCCATGCAAGACTCTGCAGACTCTCTCTCAAG	542
464	Db	rSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSerAr	484
543	QY	GAAACTCTACGTTTCCCTCTTGTGTGTATCAAAAACCTTCGGACGGAAACTGCACTTGTA	602
484	Db	gAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeuTy	504
603	QY	TTCCCATCCATCATCTGGGCTTCGCAAGATTCTCTATGGGAGTGGGCTCAGTCCGTT	662
504	Db	rSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPh	524
663	QY	TCCTCTGGCTCAGTTTACTAGTGCCTTTCTTCAGTGGTTCTGATGGGCTTTCCCCACTG	722
524	Db	eLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHisCy	544
723	QY	TTTGGCTTTCAGTTATATGATGATGTGTATTTGGGGCGAAGTCTGTACAACATCTTGA	782
544	Db	sLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGl	564
783	QY	GTCCCTTTTACCTCTATTACCAATTTCTTTGTCTTTGGGTATACATTTAAACCCCTAA	842
564	Db	uSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAs	584
843	QY	TAAACCAACAGTTGGGGCTACTCCCTTAACTTTCATGGGATATGTAATTCGGAAGTTGGG	902
584	Db	nLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrpGl	604
903	QY	TACTTTTACCGCAGGAACATATTGACTTAAACACTCAAGCAATGTTTCGAAAACTGCCTGT	962
604	Db	yThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuProVa	624
963	QY	AAATAGACCTATTGATTGGAAGTATGTCAAAGAAATGTGGGTCTTTTGGGCTTTGTGTC	1022
624	Db	lAsnSerProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAl	644
1023	QY	CCCTTTTACACAATGTGGCTATCCTGCCTGTATGTCCTTATATGTCATGTATACAATCTAA	1082
644	Db	aProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerLy	664
1083	QY	GCAGGCTTTTCACTTTCTCGCCAACTTACAAGGCTTTTCTGTGTAAA CAATATCTGAACCT	1142
664	Db	gGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsnLe	684
1143	QY	TTTACCGGTTGCCGGCAACGGTCCGGTCTCTGCGCAAGTGTGTGCTGACCGAACCCCGAC	1202
684	Db	uTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProTh	704
1203	QY	TGGATGGGCTTCGGCCATAGGCCATCAGCCCATGGCTGGAACTTTTCTGGCTCTCTGCC	1262
704	Db	rgLlyTrpGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLeuPr	724
1263	QY	GATCCCATACTCGCGAACTCTAGCAGCTTGTTTTGTCTCGACGCGGGTCTCTGAGCAAAACT	1322

Db 724 olleHlsThAlaGluLeuAlaAaCyPheAlaArgSerArgSerGlyAlaLysLe 744
Qy 1323 TATCGGAACGACAACTCTGTGCTCTCTCGGAAATACACCTCTTCATGGCTGCT 1382
Db 744 uileGlyThrAspAsnSerValValLeuSerArgLysThrSerPheProTrpLeuLe 764
Qy 1383 AGGTGTGCTGCCAACTGATCTCGCGGACGCTCTTTGTCTACGTCCTCGGCGCT 1442
Db 764 uGlyCyAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTrpValProSerAlaLe 784
Qy 1443 GAATCCCGGACGACCGCTCTCGGCGCGCTTTGGGCTCTACCGTCCCTTCTTCATCT 1502
Db 784 uAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTrpArgProLeuLeuLe 804
Qy 1503 GCCGTTCCGGCGACACGCGGCGCAGCTCTCTTTAGCGGCTCTCCCGCATGTGCGCTTC 1562
Db 804 uProPheArgProThrThrGlyArgThrSerLeuTrpAlaValSerProSerValProSe 824
Qy 1563 TCATCTCGCGACGCTGTGCACTTCGCTTCACTCTGACGTCGACGTCGAGACCGC 1620
Db 824 rHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaItrpArgProPro 843

RESULT 4
US-08-591-502B-59
; Sequence 59, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESS: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 843 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 59:
US-08-591-502B-59

Alignment Scores:			
Se-248	Length:	843	
2807.00	Matches:	527	
Percent Similarity:	Conservative:	5	
Best Local Similarity:	Mismatches:	8	
Query Match:	Indels:	2	
DB:	Gaps:	0	
US-10-761-006A-1 (1-3215) x US-08-591-502B-59 (1-843)			
QY	1	CTCCACACATTCCACCAAGCTCTCTAGATCCAGGGTGAGGGGCTATATTTCTCTGC	60
DB	304	LeuHisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys	323
QY	61	TGCTGGCTCCAGTTCCGGAACAGTAAACCTGTTCCGACTACTGCTCTCTCCATATATGC	120
DB	324	TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal	343
QY	121	AATCTTCTCGAGACTGGGGACCTCGACCGAACTGGAGAACACACATCAGGATTCCT	180
DB	344	AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro	363
QY	181	AGGACCTCTGCTGTGTACAGCGGGGTTTTCTCGTTGACAAAGAAATCTCTCAATACC	240
DB	364	ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr	383
QY	241	GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGAGCAGCCACGTGTC	299
DB	384	ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se	403
QY	300	CTGGGCCAAATTCGACGTCCCAACCTCCAAATCCTCACCAACCTCTGTCTCTCAATTT	359
DB	403	rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe	423
QY	360	GTCTGTGCTATCGCTGGATGTGTCTGCGGGTTTTTATCATATTTCTCTTCTCTGTC	419
DB	423	uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl	443
QY	420	TATGCTCATCTTCTGTGTTCTCTCGACTACCAAGGTATGTTCCTGCTCTCTCTCTC	479
DB	443	aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe	463
QY	480	TACTTCCAGGAACATCAACACGACGCGGGCCATCAAGACCTCGACGACTCTCTGCTC	539
DB	463	rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe	483
QY	540	AAGGAAACTCTACGTTTCCCTCTTGTGTGTACAAACCTTCGAGCGGAAACTGCAC	599
DB	483	rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe	503
QY	600	GTATTCCTCCATCCATCCTCGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC	659
DB	503	uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr	523
QY	660	GTCTCTCTGGCTCAGTTTACTAGTGCATTTCTTCAGTGTGTTCTGTTAGGGCTTTCC	719
DB	523	oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi	543
QY	720	CTGTTTGGCTTTTCAAGTTATATGGATATGGTATTTGGGGCGGAGTCTGTACAACATCT	779
DB	543	sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe	563
QY	780	TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTTGGGTATACATTTAAACCC	839
DB	563	uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr	583
QY	840	TAATAAAACCAACGTTGGGGCTACTCCCTTTAACTTTCATGGGATATGTAATTTGAAGTTG	899
DB	583	oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr	603
QY	900	GGGTACTTTTACCGCAGGAACATATGTGTACTTAAACTCAAGCAATGTTTTTGAAGAACTGCC	959

603 pGlyThrLeuProGlnGluHisIleValGlnLysLeuLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAAATAGACCTATTGATTGGAAGATATCTCAAAGATTGTGGCTTTTGGCGCTTTC 1019
Db 623 oValaenArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGCCCTTTTACACAATGTGGCTATCTGCTGCTGATGCTCTTTATATGATGATATACAATC 1079
Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
QY 1080 TAAGCAGGCTTTTCACTTTCTCGCAACTCTACAAAGGCTTTCTGTGTAACAATATCTGAA 1139
Db 663 xLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 683
QY 1140 CTTTACCCTGTTGCGCGCAAGCTCGGTCTCTGCGCAAGTGTGCTGACCGCAACCCC 1199
Db 683 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
QY 1200 CACTGGATGGGCTTGCCCATAGCCATCAGCGCATGGCTGGAACCTTTCTGGCTCTCTCT 1259
Db 703 oThrGlyTrpGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe 723
QY 1260 GCGATCTACTACGCGAACTCTAGCAGCTTGTGTCGCGAGCGGCTCTCGAGCAAA 1319
Db 723 uProIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLyl 743
QY 1320 ACTTATCGGAACGACACTCTGTTGCTCTCTCGGAATACACCTCTCTTCCATGGCT 1379
Db 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
QY 1380 GCTAGGTGTGTCGCAACTGATCTCTGCGCGGAGCTCTCTGCTAGTCCCGTCGCGC 1439
Db 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
QY 1440 GCTGAATCCCGCGGACGACCGCTCTCGGGCGCTTTGGGGCTCTACCGTCCCTTCTTCA 1499
Db 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
QY 1500 TCTGCGGTTCCGCGGACGACGCGGCGACCTCTCTTTAGCGGCTCTCCCGTATGTGC 1559
Db 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
QY 1560 TTTCTATCTGCGCGGCGGTGTGACCTTCGCTTCACTCTGCGGTGCGCATGGAGACCC 1619
Db 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaAlaTrpArgProPr 843
QY 1620 G 1620
Db 843 O 843

RESULT 5

US-08-591-502B-47

; Sequence 47, Application US/08591502B

; Patent No. 6607727

; GENERAL INFORMATION:

; APPLICANT: Chisari, Francis V.

; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T

; Lymphocyte Responses to Hepatitis B Virus

; NUMBER OF SEQUENCES: 99

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/591,502B

; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 843 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 47:
US-08-591-502B-47

Alignment Scores:
Pred. No.: 2,75e-246 Length: 843
Score: 2788.00 Matches: 523
Percent Similarity: 97.97% Conservative: 7
Best Local Similarity: 96.67% Mismatches: 10
Query Match: 47.20% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x US-08-591-502B-47 (1-843)

QY 1 CTCACAACATTCACCAAGCTCTGTAGATCCAGGGTGAGGGGCTATATTTCTCTGC 60
Db 304 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 323
QY 61 TCGTGGCTCCAGTTCGGGACAGTAAACCTCTCGGACTAGTCTCTCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
QY 121 AATCTTCTCGAGACTGGGGACCTCGCACCAACATGGAGAACACAAACATCAGGATTCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGGACCCCTGCTCGTGTATACAGCGGGGTTTTTCTCGTTGACAAAGATCCTCACATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 GCAGAGTCTAGACTCTG-GTGGACTCTCTCAATTTCTAGGGGAGGACCCACCTGTTTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
QY 300 CTGGGCAAAATTCGACGTCCCAACCTCCAATCACTCACCAACCTCTTCTCTCCAATTT 359
Db 403 TrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
QY 360 GTCTGGCTATCGCTGGATGTCTTCGCGCGGTTTTATCATATTCCTTTCATCTCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
QY 420 TATGCTCATCTCTTGTGTTCTCTCGACTACCAAGGTATGTTGCCGTTGCTCTCTC 479
Db 443 aMetProHisLeuLeuValIglySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
QY 480 TACTTCCAGGAACATCAACACACGACGCGGCGCCATGCAAGACCTGCACGACTCTCTGCTC 539
Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483

QY 540 AAGGAACTCTACCTTTCCCTCTGTTGCTGTACAAAACCTTCGACGGAAACTGCACATT 599
Db |||||
QY 483 rArgAsnLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
Db |||||
QY 600 GTATTCCCATCCCATCTCGGCTTTCCGACGATTCTCTATGGAGTGGGCTTCAGTCC 659
Db |||||
QY 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Db |||||
QY 660 GTTTCCTCGCTCAGTTACTAGTGCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCCA 719
Db |||||
QY 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
Db |||||
QY 720 CTGTTGGCTTCAGTTATATGATATGATATGGGGCGGAGTCTCTACACATCT 779
Db |||||
QY 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
Db |||||
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTGTTGCTTGGGTATACATTTAAACCC 839
Db |||||
QY 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
Db |||||
QY 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACCTTCATGGATATGTAATTTGGAATTG 899
Db |||||
QY 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
Db |||||
QY 900 GGGTACTTTACCGAGGACATATGTACTAACTCAAGCAATGTTTCGAAACTGCC 959
Db |||||
QY 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
Db |||||
QY 960 TGTAATAGACTATTGATTGGAAGTATGTCAAGAAATCTGGGTCTTTGGGCTTTGC 1019
Db |||||
QY 623 oValAsnArgProIleAspTrpLysValCysGluArgIleValGlyLeuLeuGlyPheAl 643
Db |||||
QY 1020 TGCCCCCTTTACAAATGTGCTATCTCGCTTGTATGCTTTATATGATGATATACAAATC 1079
Db |||||
QY 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
Db |||||
QY 1080 TAAGAGGCTTTCACTTTCTCGCAACTTACAGGCTTTCTGTGTAAACAATATCTGAA 1139
Db |||||
QY 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysGlnGlnTyrLeuHi 683
Db |||||
QY 1140 CCTTTACCGCTGTCGGGCAACGCTCGGCTCTCTCGCAAGTGTGTCGACGCAACCCC 1199
Db |||||
QY 683 sLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
Db |||||
QY 1200 CACTGGATGGGCTTGGCCATAGGCCATCAGCGCATGCTGGAACCTTTCTGGCTCTCT 1259
Db |||||
QY 703 oThrGlyTrpGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValValProLe 723
Db |||||
QY 1260 GCCGATCCATACCTGGGAACCTCTAGCAGCTGTTTCTCGACCGGCTCTGGAGCAA 1319
Db |||||
QY 723 uProIleHisThrAlaGluLeuAlaCysPheAlaAspArgSerGlyAlaLy 743
Db |||||
QY 1320 ACTTATCGGAACCGACAACTCTGTTGCTCTCTCGGAAATACACCTCTTTTCCATGGCT 1379
Db |||||
QY 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
Db |||||
QY 1380 GCTAGGGTGTGTCGCAACTGATCCTCGGGGAGCTCTTTTGTCTACGTCCCGTCGGC 1439
Db |||||
QY 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
Db |||||
QY 1440 GCTGAATCCCGGAGACCGCTCTCGGGGCGGCTTGGGCTCTACCGTCCCTCTCTTCA 1499
Db |||||
QY 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuSe 803
Db |||||
QY 1500 TCTGCGGTTCGGCGGACCGACCGGCGGACCTCTCTTTACGCGGCTCCCGTATGTGCC 1559
Db |||||
QY 803 rLeuProPheGlnProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
Db |||||
QY 1560 TTCTCATCTGCGGACCGGTGTGCATCTGCTTTCACCTCTGCACTGCGATGAGACCCACC 1619
Db |||||
QY 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTyrArgProPr 843
Db |||||

QY 1620 G 1620
Db 843 o 843
RESULT 6
US-08-591-502B-46
; Sequence 46, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 845 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 16..18
; OTHER INFORMATION: /product= "OTHER"
; /note= "Xaa = unknown"
; SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-08-591-502B-46
Alignment Scores:
Pred. No.: 6,52e-245 Length: 845
Score: 2773.00 Matches: 519
Percent Similarity: 97.04% Conservative: 6
Best Local Similarity: 95.93% Mismatches: 15
Query Match: 46.94% Indels: 2
DB: 2 Gaps: 0
US-10-761-006A-1 (1-3215) x US-08-591-502B-46 (1-845)
QY 1 CTCACAAACATTCACCAAGCTCTGCTAGATCCAGGGTCTATATTTCTCTGC 60
Db |||||
306 LeuHisAsnIleProSerSerAlaArgProGlnSerGluGlyProIleLeuSerCys 325

QY 61 TGTTGGCTCCAGTTCGGGAACAGTAACCCCTGTTCCGACTACTGCTCTCCCATATCGTC 120
 DB 326 TriPLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 345
 QY 121 AATCTTCTCGAGACTGGGGACCTCGACCGAATCGGAGAACACACATCAGGATTCCT 180
 DB 346 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyHisAsnIleArgIlePro 365
 QY 181 AGGACCCCTCTCGTGTATACAGCGGGGTTTTCTCGTTGACAAAGATCCTCAACAATACC 240
 DB 366 ArgThrProIleArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 385
 QY 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGGAGCACCCACGTTGTC 299
 DB 386 ThrGluSerThrLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 405
 QY 300 CTGGCCAAATTCGCAGTCCCGAACCTCCGAATCACTCACCACCACTCTGTCTCCCAATT 359
 DB 405 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 425
 QY 360 GTCTGGCTATCGTGTGATGTGTGCGGGGTTTTATCATATTCCTTCTCATCTCTGTCTGC 419
 DB 425 uSerTrpLeuSerLeuAspValSerAlaIaPheTyrHisIleProLeuHisProIleAl 445
 QY 420 TATGCTCTATCTTCTTGTGTTCTTCTGGACTACCAAGGTATGTTGCCGGTTTGTCTCTC 479
 DB 445 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValValCysLeuSerSe 465
 QY 480 TACTTCGAGGAATCAACACGACGCGGGCCATGCAAGACTGCAAGACTCTCTGCTCTC 539
 DB 465 rThrSerLysAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 485
 QY 540 AAGGAACTCTAGCTTTCCTCTTGTGCTGTACAAACCTTCGGACGGAAACCTGCACATT 599
 DB 485 rArgAsnLeuTyrValSerLeuPheLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 505
 QY 600 GTATTCCTCCATCCATCTCTGGGCTTTCGCAAGATTCTCTATGGAGTGGGCTCAGTCC 659
 DB 505 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 525
 QY 660 GTTTCCTGGCTCAGTTACTAGTGCATTTGTTGAGTGGTTCGATAGGCTTCCCCCA 719
 DB 525 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 545
 QY 720 CTGTTTGGCTTTCAGTTATATGATGATGTGTTATTGGGGCGAAGTCTGTCAACAATCT 779
 DB 545 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 565
 QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
 DB 565 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 585
 QY 840 TAATAAAACAAAGTTGGGCTACTCCCTTAATCTCATCGGATATGTAATTCGAAGTTG 899
 DB 585 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlyCysTr 605
 QY 900 GGGTACTTTTACCAGGAACATATTGACTAAACCTCAAGCAATGTTTTCGAAACTGCC 959
 DB 605 pGlyThrLeuProGlnGluHisIleValLeuLysIleValLeuLysGlnCysPheArgLysLeuPr 625
 QY 960 TGTAATAGACCTATTGATTGGAAGTATGTCAAAGAAATTGGGGTCTTTTGGGCTTTGC 1019
 DB 625 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 645
 QY 1020 TGCCCTTTTACAAATGTGGCTATCTGCTGTGATGCCCTTTATATGATGATATACATC 1079
 DB 645 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 665
 QY 1080 TAAGCAGCTTTCATCTTCTCGCAACTCAACAGGCCCTTTCTGTGTAAACAAATATCTGAA 1139
 DB 665 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuLeuCysLysGlnTyrLeuAs

QY 1140 CTTTATACCCCGTTGCCCGCAACGGTCCGGTCTCTGCCAAGTGTGTTGTCGACCAACCCC 1199
 DB 685 nLeuTyrProValAlaAArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 705
 QY 1200 CACTGGATCGGCTTGGCCATAGGCCATCAGCGCATGGCTGGAACTTTTCTGGCTCTCT 1259
 DB 705 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 725
 QY 1260 GCCGATCCACTACTGGCGAACTCCTAGCAGCTTGTGTTGCTCGCAGCCGGTCTGGAGCAA 1319
 DB 725 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaAArgSerArgSerGlyAlaL 745
 QY 1320 ACTTATCGGAACGCAACTCTGTCTCTCTCTCTCTCGAAATACACCTCTTCCATGGCT 1379
 DB 745 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 765
 QY 1380 GTTAGGCTGTGTCACACTCGTCGCGGACGTCCTTTTGTCTAGTCCCGTCCGC 1439
 DB 765 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 785
 QY 1440 GCTGAATCCCGCGACGACCCGCTCTCGGGCGGCTTTGGGGCTCTACCGTCCCTTCTTCA 1499
 DB 785 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 805
 QY 1500 TGTCCGTTCCGGCCGACACCGGGCGGACCTCTCTTTAGCGGCTCTCCCGTATGTGCC 1559
 DB 805 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 825
 QY 1560 TTCTCATCTCGCGGACCGGTGTCACCTTCCTTCCCTCTGCACTGCGACGTGGAGACCACC 1619
 DB 825 oSerHisLeuProAspArgValHisPheProserProLeuHisValAlaTrpArgProPr 845
 QY 1620 G 1620
 DB 845 O 845
 DB 845 O 845
 RESULT 7
 US-08-591-502B-48
 ; Sequence 48, Application US/08591502B
 ; Patent No. 6607727
 ; GENERAL INFORMATION:
 ; APPLICANT: Chisari, Francis V.
 ; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
 ; Lymphocyte Responses to Hepatitis B Virus
 ; NUMBER OF SEQUENCES: 99
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/591,502B
 ; FILING DATE: 20-May-1996
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/749,540
 ; FILING DATE: 26-AUG-1991
 ; APPLICATION NUMBER: US 07/935,898
 ; FILING DATE: 26-AUG-1992
 ; APPLICATION NUMBER: US 08/100,870
 ; FILING DATE: 02-AUG-1993
 ; APPLICATION NUMBER: WO PCT/US94/08685
 ; FILING DATE: 01-AUG-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Weber, Ellen Lauver
 ; REGISTRATION NUMBER: 32,762

REFERENCE/DOCKET NUMBER: 014740-000230US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 48:

SEQUENCE CHARACTERISTICS:

LENGTH: 843 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 48:

US-08-591-502B-48

Alignment Scores:
Pred. No.: 1.25e-243 Length: 843
Score: 2759.00 Matches: 519
Percent Similarity: 97.23% Conservative: 7
Best Local Similarity: 95.93% Mismatches: 14
Query Match: 46.71% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x US-08-591-502B-48 (1-843)

```

Qy 1 CTCCACAACTTCCACCAAGCTCTGCTAGATCCAGGCTGAGGGCCCTATATTTCTCTGC 60
Db 304 LeuHisAsnIleProProSerCysAlaArgSerGlnSerGluGlyProIleSerCys 323
Qy 61 TGTGGCTCCAGTTCGGGAACAGTAACCCCTGTTCCGACTACTCCCTCTCCCATATGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerGluProCysSerAspTyrCysLeuThrHisIleVal 343
Qy 121 AATCTTCTCGAGAGCTGGGACCTGACCCGAAATGAGGAAACACATCAGATTCCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
Qy 181 AGGACCTCTGCTGTTTACAGGGGGGTTTCTGTTGACAAAGATCCTCACAAATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspIleAsnProHisAsnThr 383
Qy 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCAATTTTCTAGGGGGAGCACCCAGTGTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
Qy 300 CTGGCAAAATTCGAGTCCCAACCTCAATCACTACCAACCTCTGCTCTCAATTT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
Qy 360 GTCTGGCTATCGTGGATGCTCTGCGCGGTTTATCATATTCCTTCTTCTCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
Qy 420 TATGCTCATCTCTCTGTTGTTCTTCTGGACTACCAAGGATGTTGCGCGTTGTCCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
Qy 480 TACTTCCAGGAACATCAACACACGACGAGGGGCCATGCAAGCTGACGACTCTGCTC 539
Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
Qy 540 AAGGAACTCTACGTTTCCCTCTGCTGCTGACAAACCTTCGGACGAAACTGCACCT 599
Db 483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisIle 503
Qy 600 GTATTCCCATCCATCATCTCGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
Db 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Qy 660 GTTCTCTCGGCTCAGTTTACTAGTCCCAATTTGTTTCAAGTGTCTAGGCTTTCCCCCA 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
Qy 720 CTGTTTGGCTTTTCAATTATGATGATGTTGTTTGGGGGCAAGTCTGTACACATCT 779

```

```

Db 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
Qy 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTCGGGTATACATTTAAACCC 839
Db 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
Qy 840 TAATAAAACCAACAGTTGGGGCTACTCCCTTAACCTTCATGGGATATGTAATTTGGAAGTTG 899
Db 583 oHisLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
Qy 900 GGCTACTTTACCGCAGGAACATATTGTACTAAACCTCAACCAATGTTTTCGAAAACCTGCC 959
Db 603 pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr 623
Qy 960 TGTAAATAGACCTATTGATTGGAAAGATGATCTCAAGAAATTTGTTGGTCTTTTGGGCTTTGC 1019
Db 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
Qy 1020 TGCCCTTTTACACAATGTGGCTATCTGCTCTGATGCTTATATCATGATATACAAATC 1079
Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
Qy 1080 TAAGACGGCTTTTCTCTTCTCGCAACTTACAGGCTTTTCTGTGTAACAAATATCTGAA 1139
Db 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi 683
Qy 1140 CTTTACCTCCCTGTCGGCGCAACGGTCTGCTGCAAGTGTTCGTGACGCAACCC 1199
Db 683 sLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
Qy 1200 CACTGGATGGGCTTGGCCATAGCCATCAGCGCATGGCTGGAACTTTTGGGCTCTCT 1259
Db 703 oThrGlyTrpGlyLeuAlaIleGlyGlnSerGlyMetArgGlyThrPheValAlaProLe 723
Qy 1260 GCCATCCATCTGCGGAACCTCTAGCAGCTTGTGTTGCTCGCAGCGGCTCTGGAGCAA 1319
Db 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy 743
Qy 1320 ACTTATCGGAACCGACAACTCTGTTGCTCTCTCGGAAATACACCTCTTTCATGGCT 1379
Db 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpue 763
Qy 1380 GCTAGGTTGCTGCTCCAACTGGATCTGCGGGGACGCTCTTGTCTACGTCCTGCGC 1439
Db 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
Qy 1440 GCTGAATCCCGGACGACCGCTCTCGGGGCGGTTTGGGGCTCTACCGTCCCTCTTCA 1499
Db 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
Qy 1500 TCTCCGTTCCGGCGGACCGGCGGCGCACCTCTCTTACCGGCTCTCCCGTATGTGCC 1559
Db 803 sLeuProPheArgProThrThrGlyArgAlaSerLeuTyrAlaValSerProSerValPr 823
Qy 1560 TTCTCATCTGCGGACCGCTGTCACCTTCCGTTTCACTCTGACGCTGCGATGAGACCAACC 1619
Db 823 oSerHisLeuProValArgValHisPheAlaSerProLeuHisValAlaIleTrpArgProPr 843
Qy 1620 G 1620
Db 843 o 843

```

RESULT 8

US-08-416-950-11

; Sequence 11, Application US/08416950

; Patent No. 5780036

; GENERAL INFORMATION:

; APPLICANT: CHISARI, Francis V.

; TITLE OF INVENTION: PEPTIDES FOR INDUCING CYTOTOXIC T

; TITLE OF INVENTION: LYMPHOCYTE RESPONSES TO HEPATITIS B VIRUS

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend Khourie and Crew
 ; STREET: One Market Plaza, Steuart Street tower
 ; CITY: San Francisco
 ; STATE: CA
 ; COUNTRY: U.S.A.
 ; ZIP: 94105-1492
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/416,950
 ; FILING DATE:
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US
 ; FILING DATE:
 ; APPLICATION NUMBER: US 07/935,898
 ; FILING DATE: 26-AUG-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/749,540
 ; FILING DATE: 26-AUG-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Pameleer, Steven W.
 ; REGISTRATION NUMBER: 31,990
 ; REFERENCE/DOCKET NUMBER: 14740-2-2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (206) 467-9600
 ; TELEFAX: (415) 543-5043
 ; INFORMATION FOR SEQ ID NO: 11:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 845 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-416-950-11

Alignment Scores:
 Pred. No.: 1,25e-243 Length: 845
 Score: 2759.00 Matches: 516
 Percent Similarity: 97.23% Conservative: 10
 Best Local Similarity: 95.38% Mismatches: 14
 Query Match: 46.71% Indels: 2
 DB: 1 Gaps: 0
 US-10-761-006a-1 (1-3215) x US-08-416-950-11 (1-845)

QY 1 CTCACAACTTCCACAGCTCTGTAGATCCAGGGTGAGGGGCTATATTTCTCTGC 60
 Db 306 LeuHisAsnIleProProAsnSerAlaArgSerGlnSerGlyProValPheSerCys 325
 QY 61 TGTGGCTCCAGTCCGGACAGTAAACCTGTTCCGACTACTGCTCTCCCATATCGTC 120
 Db 326 TrpIrpPheGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 345
 QY 121 AATCTTCTCTGAGACTGGGACCTCGACCGAATCGGAGAACACAACTCAGGATTCCCT 180
 Db 346 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 365
 QY 181 AGAACCTCTGCTGTTACAGCGGGGTTTTCTCGTTGACAAGAACTCTCAATATACC 240
 Db 366 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 385
 QY 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGAGACCCACGCTTTC 299
 Db 386 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrArgVal-Se 405
 QY 300 CTGGCCAAATTCGAGTCCCAACCTCCCAATCACTCACCACCTCTGTCTCTCCCAATTT 359
 Db 405 rTTPProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 425

QY 360 GTCTCTGCTATCGCTCGATGTGTCTCGGGGCTTTTATCATATTCCTTCTTCATCTCTGCTGC 419
 Db 425 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 445
 QY 420 TATGCTCATCTCTTCTTGTGTCTCTGACTACCAAGGTATGTGTCCCGTTGTCTCTC 479
 Db 445 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 465
 QY 480 TACTTCCAGGAACATCAACACCCAGCACGGGGCCATGCAAGACCTCGACACTCTCTGCTC 539
 Db 465 rAsnSerArgIleAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 485
 QY 540 AAGAAACTCTAGCTTCTCTCTGTGTCTGTACAAAACCTTCGGAGCGGAACCTGCACIT 599
 Db 485 rArgAsnLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 505
 QY 600 GTATTCCCATCCCATCATCTCTGGGCTTCGCAAGATTCTCTATGGAGTGGGCTCTAGTCC 659
 Db 505 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 525
 QY 660 GTTCTCTCTGGCTCAGTTTACTAGTCCCATTTTGTTCAGTGGTTCGTAGGGCTTTCCCCCA 719
 Db 525 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 545
 QY 720 CTGTTTGGCTTTCAGTTATATGATGATGTGTGTATTTGGGGGGGGAAGTCTGTACAACTCT 779
 Db 545 sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe 565
 QY 780 TGAAGTCCCTTTTACCTCTTACCAATTTCTCTTGTCTTGTGGGTATACATTTAAACCC 839
 Db 565 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 585
 QY 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACTTCATGGGATATGTAAATCGAAGTTG 899
 Db 585 oAsnLysThrLysArgTyrGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 605
 QY 900 GGGTACTTTACGCGAGGAACATATGTACTTAAACTCAAGCAATGTTTTCGAAACTGCC 959
 Db 605 pGlyThrIleProGlnGluHisIleValGlnLysIleLysGlnCysPheArgLysLeuPr 625
 QY 960 TGTAAATAGACCTATTGATTGGAAGTATGTCAAGAAATGTGGGTCTTTTGGGCTTTGC 1019
 Db 625 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 645
 QY 1020 TGCCCTTTTACAAATGTGGCTATCTCTGCTTATGCTCTTTATATGCATGTATACAATC 1079
 Db 645 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 665
 QY 1080 TAAGCAGGCTTTCACCTTCTCGCAACTTACAAGGCTTCTGTGTAAACAATATCTGAA 1139
 Db 665 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 685
 QY 1140 CCTTTACCCCGTTTGGCCGCAACGGTCTCTCGCAAGTGTCTTGTGACCAACACCCC 1199
 Db 685 nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 705
 QY 1200 CACTGATCGGGCTTGGCCATAGGCCATCAGCGCATGGCTGGAACTTTCTGCTCTCTCT 1259
 Db 705 oThrGlyTyrGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 725
 QY 1260 GCGGATCCATCTCGGGAACCTCTAGCAGCTTGTGTGTCTCGCAGCGGCTCTGGAGCAA 1319
 Db 725 uProIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLy 745
 QY 1320 ACTTATCGGAACCGCAACTCTGTCTCTCTCGGAATATACACCTCTCTTTCATGGCT 1379
 Db 745 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 765
 QY 1380 GTTAGGGTGTGTGCAACTGGAATCTGTGCGGGAGCTCTTTGTCTAGCTCCGCTGGC 1439
 Db 765 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 785
 QY 1440 GCTGAATCCCGGGAGCAGCCCGTCTCGGGGCGGTTTGGGGCTCTACCGTCCCTTCTTCA 1499

Db 785 aLeuAsnProHlaAspProSerArgGlyArgLeuGlyLeuTyArgProLeuLeuAr 805
Qy 1500 TCTCCGCTCCGGCCGACCGGGCGACCTCTCTTTACGGGTCTCCCGGTATGTGCC 1559
Db 805 gLeuProPheArgProThrThrGlyA-gThrSerLeuTyAlaValSerProSerValPr 825
Qy 1560 TTCTCANTCCGACCGGTGGCACTTCGCTTCACTCTCGACGTCGACATGGAGACCA 1619
Db 825 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaItrArgProPr 845
Qy 1620 G 1620
Db 845 o 845

RESULT 9

US-08-469-830-11
; Sequence 11, Application US/08469830
; Patent No. 5932224
; GENERAL INFORMATION:
; APPLICANT: CHISARI, Francis V.
; TITLE OF INVENTION: PEPTIDES FOR INDUCING CYTOTOXIC T
; TITLE OF INVENTION: LYMPHOCYTE RESPONSES TO HEPATITIS B VIRUS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94105-1492
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,830
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14740-2-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 845 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-469-830-11

Alignment Scores:

Pred. No.: 1-25e-243 Length: 845
Score: 2759.00 Matches: 516
Percent Similarity: 97.23% Conservative: 10
Best Local Similarity: 95.38% Mismatches: 14
Query Match: 46.71% Indels: 2
DB: 1 Gaps: 0

US-10-761-006A-1 (1-3215) x US-08-469-830-11 (1-845)

Qy 1 CTCACAAACATTCACCAAGCTCTGCTAGATCCAGGGTGAGGGGCTATATATTTCTCTGC 60
Db 306 LeuHisAsnIleProProAsnSerAlaArgSerGlnSerGluGlyProValPheSerCys 325
Qy 61 TGTGGCTCAGTTCGGGAACAGTAACCTGTTCCGACTACTGCTCTCCCATATGCTC 120
Db 326 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrcysLeuThrHisIleVal 345
Qy 121 AATCTTCTCGAGACTGGGGACCTCCACGAAATGAGAAACACAAACATCAGGATTCCT 180
Db 346 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyHisAsnIleArgIlePro 365
Qy 181 AGGACCCCTCTCTGTTTACAGCGGGGTTTTCTCTGTTGACAAGATCTCTCAATACC 240
Db 366 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 385
Qy 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGGAGCACCCAGCTGTC 299
Db 386 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrArgVal-Se 405
Qy 300 CTGGCCAAAATTTCGAGTCCCAACCTCCCAATCACCACCACTCTGTCTCTCAATTT 359
Db 405 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 425
Qy 360 GTCCTGGCTATCGCTGGATGTCTCGCGGGTTTTATCATATTCCTTCTCATCTCTGCTC 419
Db 425 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyzHisIleProLeuHisProAlaAl 445
Qy 420 TATGCTCATCTTCTGTTGTTCTCTGACTACCAAGGTATGTTGCCGCTTTGCTCTC 479
Db 445 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyzValAlaArgLeuSerSe 465
Qy 480 TACTTCCAGGAACATCAACACGACGCGGGCCATCGAAGACCTGCACGACTCTCTGCTC 539
Db 465 rAsnSerArgIleAsnTyzGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 485
Qy 540 AAGGAAACTCTACGTTTCCCTCTTGTGTGTACAAAACCTTCGGAGCGGAACTGCACCT 599
Db 485 rArgAsnLeuTyzValSerLeuLeuLeuLeuTyzLysThrPheGlyArgLysLeuHisLe 505
Qy 600 GTATTCCCATCCCATCATCTCGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
Db 505 uTyzSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 525
Qy 660 GTTTCCTCTGCTCAGTTTACTAGTGCATTTGTTTCAGTGTCTCGTAGGGCTTTCCCA 719
Db 525 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProH 545
Qy 720 CTGTTTGGCTTTTACCTTATACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
Db 545 sCysLeuAlaPheSerTyzMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 565
Qy 780 TGAGTCCCTTTTACCTTATACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
Db 565 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 585
Qy 840 TAATAAACAACGTTGGGCTACTCCCTTAACCTTCATGCGATATGTAATTTGGAAGTTG 899
Db 585 oAsnLysThrLysArgTrpGlyTyzSerLeuAsnPheMetGlyTyzValIleGlySerTr 605
Qy 900 GGTTACTTTTACCGAGGAACATATTGTACTAAAACTCAAGCAATGTTTTTCGAAAACTGCC 959
Db 605 pGlyThrIleProGlnGluHisIleValGlnLysIleLysGlnCysPheArgLysLeuPr 625
Qy 960 TGTAAATAGACCTATTGATTGGAAAGTATGTCAAAGAAATTTGGGTCTTTTGGGCTTTGC 1019
Db 625 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 645
Qy 1020 TGCCCCCTTTTACAAATGTGGCTATCTCGCTTGTGATGCTTTATATGATGATATACAATC 1079
Db 645 aAlaProPheThrGlnCysGlyTyzProAlaLeuMetProLeuTyzAlaCysIleGlnSe 665

```
QY 1080 TAAGCAGGCTTTACATTTCTCGCAACTACAGGCGCTTTCTGTGTAAACAATATCTGAA 1139
Db 665 rLysGlnAlaPheThrPheSerProThrTyrlYsAlaPheLeuCysLysGlnTyrlLeuAs 685
QY 1140 CTTTACCCCGTTGCCGCAACGGTCCGCTCTCCCAAGTGTTCGTGACGCAACCCC 1199
Db 685 nLeuTyrlProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 705
QY 1200 CACTGGATGGGCTTGCCATAGCCATCAGCCATGGCTGGAACTTTCTGCTCTCTCT 1259
Db 705 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 725
QY 1260 GCCGATCCATCTCGGAACTCTAGCAGCTTGTTCGTCGAGCGGTCTGGAGCAAA 1319
Db 725 uProIleHisThrAlaGluLeuLeuAlaLaCysPheAlaArgSerArgSerGlyAlaLy 745
QY 1320 ACTTATCGGAACCGACAACTCTGTGCTCTCTCGGAATACACCTCTTTCATGGCT 1379
Db 745 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrlThrSerPheProIrlPle 765
QY 1380 GCTAGGCTGTGCTGCCAACTGGATCTCTGCGGGAGCTCTCTTTGTCTACGTCGCCGTCGCC 1439
Db 765 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrlValProSerAl 785
QY 1440 GCTGAATCCCGCGGACGCCGCTCTGGGGCGCTTGGGGCTCTACGTCCTCTTCTTCA 1499
Db 785 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrlArgProLeuLeuAr 805
QY 1500 TCTGCGCTTCCGCGCACACCGCGGCACCTCTCTTTACGCGGTCTCCCGATGTGCC 1559
Db 805 gLeuProPheArgProThrThrGlyArgThrSerLeuTyrlAlaValSerProSerValPr 825
QY 1560 TTCTCATCTGCCGACCGTGTGCACTTCTGCTTCCACTCTGCACTGCGAGTCGAGACCAACC 1619
Db 825 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 845
QY 1620 G 1620
Db 845 o 845

RESULT 10
US-08-591-502B-11
; Sequence 11, Application US/08591502B
; Patent No. 660727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatenIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
```

```
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 845 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..845
; OTHER INFORMATION: /product= "OTHER"
; /note= "Xaa = any amino acid
; (-50% consensus)"
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-08-591-502B-11

Alignment Scores:
Pred. No.: 5,566-242 Length: 845
Score: 2741.00 Matches: 514
Percent Similarity: 96.30% Conservative: 7
Best Local Similarity: 95.01% Mismatches: 19
Query Match: 46.40% Indels: 2
DB: Gaps: 0

US-10-761-006A-1 (1-3215) x US-08-591-502B-11 (1-845)
QY 1 CTCACAACATTTCCACCAAGCTCTGCTAGATCCAGGGTGAGGGGCTATATTTTCTCTGC 60
Db 306 LeuHisAsn***ProProAsnSerAlaArgSerGlnSerGlnSerGlyProValPheSerCys 325
QY 61 TGGTGGCTCCAGTTCCGGAACAGTAAACCCCTGTTCCGACTACTGCTCTCTCCCATATCGTC 120
Db 326 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyCysLeu**HisIleVal 345
QY 121 ATCTTCTCAGAGACTGGGGAGCCCTGCACCGAACATGGAGAACACAAATCAGGATTCCT 180
Db 346 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHis***IleArgIlePro 365
QY 181 AGACCCCTGCTCGTGTACAGCGGGGTTTTCTCGTTGACAAGAATCTCAACAATACC 240
Db 366 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 385
QY 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGGAGCACCCACGTGTTTC 299
Db 386 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGly**ThrArgVal-Se 405
QY 300 CTGGCCAAAATTCCGAGTCCCAACCTCCAATCACTCACCAACCTCTGCTCTCCCAATT 359
Db 405 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 425
QY 360 GTCTCGGCTATCGCTGGATGTCTGCGCGGTTTTATCATATTTCTCTTCTCATCTCTGCTGC 419
Db 425 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrlHisIleProLeuHisProAlaAl 445
QY 420 TATGCTCATCTTCTTGTGTTCTTCTGAGCTTACCAAGGTATGTTGCCGTTGTGCTCTC 479
Db 445 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrlValAlaArgLeuSerSe 465
QY 480 TACTTCCAGGAACATCAACCAACGAGCGGGCCATGCAAGACCTGCACGACTCTCTGCTC 539
Db 465 rAsnSerArgIleIleAsn***GlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 485
QY 540 AAGGAAACTCTACGTTTCCCTCTTGTGTGTGTACAAAACCTTCGGACGGAACCTGACACT 599
Db 485 rArgAsnLeuTyrlValSerLeuLeuLeuLeuTyrlsThrPheGlyArgLysLeuHisLe 505
```



```
Qy 600 GTATTCCCATCCATCCTGGGCTTTCCAGATTCTCTATGGAGTGGGCTCAGTCC 659
Db 505 uTySerHisProIleIleuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 525
Qy 660 GTTTCCTCTGGCTCAGTTTACTAGTCGCAATTTGTTCAGTGTCTCGTAGGGCTTTCCGCCA 719
Db 525 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 545
Qy 720 CTGTTTGGCTTTCAGTTATATGAGATGATGTGATTTGGGGCGAAGTCTGTACACATCT 779
Db 545 sCysLeuAlaPheSerTyrrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 565
Qy 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
Db 565 uGluSerLeuPheThrAlaValThrAsnPheLeuSerLeuGlyIleHisLeuAsnPr 585
Qy 840 TAAATAAACCAAGCTTGGGCTACTCCCTTAACCTTCATGGATATGTAATTTGGAATTG 899
Db 585 oAsnLysThrLysArgTrpGlyTyrrSerLeuAsnPheMetGlyTyrrValIleGlySerTr 605
Qy 900 GGGTACTTTACCGAGGAACATATTTGTAATAAACTCAGCAATGTTTTCGAAAACCTGCC 959
Db 605 pGlyThrLeuProGlnGluHisIleValGlnLysIleLysGlnCysPheArgLysLeuPr 625
Qy 960 TGTAAATAGACTATTGATTGGAAGATGTCGAAGATTGCTGGGCTCTTTGGGCTTTGC 1019
Db 625 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 645
Qy 1020 TGCCCTTTTACAAATGTGCTATCTGCTGCTGATGCTTATATGATGATATACAAATC 1079
Db 645 aAlaProPheThrGlnCysGlyTyrrProAlaLeuMetProLeuTyrrAlaCysIleGlnSe 665
Qy 1080 TAAGCAGGCTTTTCACTTTCTCGCCAACTTACAAGGCTTTCTGTGTAAACAATATCTGAA 1139
Db 665 rLysGlnAlaPheThrPheSerProThrTyrrLysAlaPheLeuCysLysGlnTyrrLeuAs 685
Qy 1140 CCTTTACCCGCTGGCCGGCAACCGTCCGCTCTGCGCAAGTGTTCGTGACGCAACCCC 1199
Db 685 nLeuTyrrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 705
Qy 1200 CACTGATGGGCTTGGGCATAGGCATCAGCGATGCTCGAACCTTTCTGGCTCTCTCT 1259
Db 705 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 725
Qy 1260 GCCGATCCATACGCGGAACCTCTAGCAGCTGTGTTTCTCGCAGCCGGTCTGAGCAAA 1319
Db 725 uProIleHisThrAlaGluLeuLeuAlaAAsPheAlaArgSerArgSerGlyAlaLy 745
Qy 1320 ACTTATCGGAACCGCAACTCTGTGCTCTCTCGGAATACACCTCTTCCATGGCT 1379
Db 745 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrrThrSerPheProTrpLe 765
Qy 1380 GCTAGGGTGTGTCGCAACTGATCTCGCGGGAGCTCTTTGTCTACGTCGCGTCCGC 1439
Db 765 uLeuGlyCysAlaAlaAsnTrpIleuArgGlyThrSerPheValTyrrValProSerAl 785
Qy 1440 GCTGAATCCCGGAGCAGCCGCTCTCGGGGCGGTTTGGGGCTCTACCGTCCCTTCTTCA 1499
Db 785 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrrArgProLeuLeuAr 805
Qy 1500 TCTGCGGTTCGCGGCGACACGCGGGCGGACCTCTCTTTACCGGGTCTCCCGTATGTGCC 1559
Db 805 gLeuProPheArgProThrThrGlyArgThrSerLeuTyrrAlaValSerProSerValPr 825
Qy 1560 TTCTCATCTGCGGACCGTGTGCTCTGCTTCACTCTGACGCTGCGATGAGACACACC 1619
Db 825 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 845
Qy 1620 G 1620
Db 845 o 845
```

```
RESULT 11
US-08-591-502B-51
; Sequence 51, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 842 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 51:
US-08-591-502B-51
Alignment Scores:
Pred. No.: 2,19e-241 Length: 842
Score: 2734.50 Matches: 517
Percent Similarity: 96.86% Conservative: 7
Best Local Similarity: 95.56% Mismatches: 15
Query Match: 46.29% Indels: 3
DB: 2 Gaps: 1
US-10-761-006A-1 (1-3215) x US-08-591-502B-51 (1-842)
Qy 1 CTCCACACATTCACCAAGCTCTGTAGATCCAGGGTGGGGCCCTATATTTTCCTGC 60
Db 304 LeuHisIleSerProSerProAlaArgSerGlnSerGluGlyProllePheSerSer 323
Qy 61 TGTGGCTCCAGTTTCCGAAACAGTAAACCCCTGTTCCGACTACTCCCTCTCCCATATCCTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrrCysLeuThrHisIleVal 343
Qy 121 AATCTTCTCAGAGGACTGGGACCTCGACCGAACATGAGAACACACATCAGATTCTCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
```

181 AGGACCCCTCGTGTGTTACAGCGGGGTTTTCTCGTTGACAGAATCCTCACAATACC 240
Db |||||
364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY |||||
241 GCAGAGTCTAGACTCTG-GTGGACTCTCTCAATTTTCTAGGGGGAGCACCCACGTGTTTC 299
Db |||||
384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
QY |||||
300 CTGGCCAAATTCGAGTCCGCCAACCTCCAACTCACAATCCTCTGTCCTCCAAATTT 359
Db |||||
403 rTTPProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 423
QY |||||
360 GTCTCGCATCGTGAATCGTCTGGCGGTTTTATCATATTCCTCTTCATCTCTGCTGC 419
Db |||||
423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
QY |||||
420 TATGCTCTCATCTCTCTGTTGTTGTTCTCTGACTACCAAGGTATGTTGCCCGTTGCTTC 479
Db |||||
443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
QY |||||
480 TACTTCAGGAACATCAACCAACGACGACGGGCCATGCAAGACTGACAGCTGCTGCTC 539
Db |||||
463 rThrSerArgAsnIleAsnIleGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
QY |||||
540 AAGGAACCTCTACGTTTCCCTCTGTTGCTGTACAAACCTTCGGACGGAAACCTGCATCT 599
Db |||||
483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
QY |||||
600 GTATTCCTCATCCCATCTCTGGGCTTTCCGAAGATTCTCATGGAGTGGCGCTCAGTCC 659
Db |||||
503 uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyGlyLeuSerPr 523
QY |||||
660 GTTTCCTCGCTCAGTTACTAGTGCATTTGTTAGTGGTTCGATAGGCTTTCCCCCA 719
Db |||||
523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
QY |||||
720 CTGTTTGGCTTTCAGTTATATGATGATGTTGTTGTTGGGGCGAAGTCTGTACAATCT 779
Db |||||
543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY |||||
780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTCTGTTGGGTATACATTTAAACCC 839
Db |||||
563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
QY |||||
840 TAATAAACCAAGTTGGGGCTACTCCCTTAATCTCATGGATATGTAATTAAGAGTTG 899
Db |||||
583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
QY |||||
900 GGGTACTTTACCGCAGGAACATATTGTACTAAACTCAAGCAATGTTTTCGMAAAGTCC 959
Db |||||
603 pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr 623
QY |||||
960 TGTAATATAGACCTATTGATTGGAAAGTATGTCAAGAAATTTGGGTCTTTGGGCTTTC 1019
Db |||||
623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY |||||
1020 TGCCCCCTTTTACAAATGTGGCTATCTCGCTTATGTCCTTATGATGCTTTATGATCATATACAATC 1079
Db |||||
643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
QY |||||
1080 TAAGCAGCTTTCATTTTCTCGCAACTTACAGGCCCTTTCTGTGTAAACAATATCTGAA 1139
Db |||||
663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi 683
QY |||||
1140 CTTTACCCGCTTGGCCGCAACGGTCTGCTCTGCAAGTGTCTGCTGACGCAACCCC 1199
Db |||||
683 sLeuTyrProValAlaArg---ArgThrAlaLeuCysGlnValPheAlaAspAlaThrPr 702
QY |||||
1200 CACTGGATGGGGCTTGGCCATAGCCATACCGCATGGCTGGAACTTTCTGGCTCTCTCT 1259
Db |||||
702 oThrGlyTrpGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe 722
QY |||||
1260 GCCGATCCATACCTCGCGAACTCTAGCAGCTTGTTTGTCTCGCAGCGGCTCTGGAGCAA 1319

Db |||||
722 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 742
QY |||||
1320 ACTTATCGAAACCGAACACTCTGTTGTCTCTCTCGGAAATACACTCTCTTTCCATGGCT 1379
Db |||||
742 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 762
QY |||||
1380 GGTAGGTGTGTGCGCAACTGGATCTCGCGGGAGCTCTTTGTCTAGTCCCGTCCGCG 1439
Db |||||
762 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrTyrPheValTyrValProSerAl 782
QY |||||
1440 GGTGAATCCCGCGGACGACCCGCTCTCGGGCGGTTTGGGGCTCTACCGTCCCCCTTTCA 1499
Db |||||
782 aLeuAsnProAlaAspAspProSerArgLysArgLeuGlyLeuIleArgProLeuLeuHi 802
QY |||||
1500 TCTGCGCTTCCGGCGACACCGGGCGACCTCTCTTTTACGGCGGTCTCCCGTATGTGCC 1559
Db |||||
802 sLeuArgPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 822
QY |||||
1560 TTCTCATCTCGCGGACCGTGTGCATCTCGCTTACCTCTGACGTCGCGATGGAGACCACC 1619
Db |||||
822 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 842
QY |||||
1620 G 1620
Db |||||
842 o 842
RESULT 12
US-08-591-502B-50
; Sequence 50, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 842 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>

```

;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 50:
US-08-591-502B-50

Alignment Scores:
Pred. No.: 6, 28e-241 Length: 842
Score: 2729.50 Matches: 516
Percent Similarity: 96.67% Conservative: 7
Best Local Similarity: 95.38% Mismatches: 16
Query Match: 46.21% Indels: 3
DB: 2 Gaps: 1

US-10-761-006a-1 (1-3215) x US-08-591-502B-50 (1-842)
QY 1 CTCACAACTTCCACCAAGCTCTGCTAGATCCAGGGTGGGGCCCTATATTTCTCTGC 60
DB 304 LeuHisIleSerProSerProAlaArgSerGlnSerGluGlyProIlePheSerSer 323
QY 61 TGTGGCTCCAGTTCGGGAACAGTAACCCCTGTTCCGACTACTCCCTCCCATATCGTC 120
DB 324 TrpTrpLeuGlnPheArgAsnSerLysProCysCysAspTyrCysLeuThrHisIleVal 343
QY 121 AATCTTCTCAGGACTGGGACCTCGACCGAATCGAGAACACACATCAGGATTCCT 180
DB 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGGACCCCTGCTGCTGTGTACAGGGGGGTTTTCTCGTTGACAAAGATCTCTCAATACC 240
DB 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 GCAGAGTCTAGACTCTGCTGCTCAATTTCTAGGGGAGCACCACGCTGTC 299
DB 384 ThrGluSerArgLeuValAlaAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
QY 300 CTGGCCAAATTCGAGTCCCACTCACTCACTCACTCACTCACTCACTCACTCACT 359
DB 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
QY 360 GTCCTGGCTATCGTGGATGTCTGCGCGGTTTTATCATATTTCTCTTCACTCTGCTGC 419
DB 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
QY 420 TATGCTCATCTCTGTTGGTCTCTGGAATCAAGAGTATGTTGCGCGGTTGCTCTC 479
DB 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
QY 480 TACTTCCAGGACATCAACCAACCCAGCAGCGGCCATGCAAGCTGACGACTCTCTGCTC 539
DB 463 rThrSerArgAsnIleAsnHisGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
QY 540 AAGGAACTCTACGTTTCCCTCTCTGCTGTACAAAACCTTCGGACGGAACTGCATCT 599
DB 483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgGlyLeuHisLe 503
QY 600 GPATTTCCCATCCATCTCGGCTTTCGCAAGATTTCTTATGGAGTGGGCTCAGTCC 659
DB 503 uTyrSerHisProIleLeuGlyPheArgLysIleProMetGlyGlyGlyLeuSerPr 523
QY 660 GTTTCCTCGGCTCAGTTTACTAGTGCATTTGTTTCAGTGGTTCTGAGGGCTTTCCCCCA 719
DB 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHI 543
QY 720 CTGTTGGCTTTTCAGTTATATGATGATGTGTTATTTGGGGGCGAAGTCTGTACACATCT 779
DB 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTTGGGTATACATTTAAACC 839
DB 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeAsnPr 583
QY 840 TAAATAAACCAACGTTGGGGCTACTCCCTTAACCTCATGGATATGTAATTTGGAAGTTG 899
  
```

```

DB 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
QY 900 GGCTACTTTTACCGCAGCAATATTTACTATAAATCTCAAGCAATGTTTTCGAAAACTGCC 959
DB 603 pGlyThrLeuProGlnGluHisIleValLeuLysIleValGlnCysPheArgLysLeuPr 623
QY 960 TGTAAATAGACCTATTGATTGAAAGATGTCAAGAAATTTGGGCTCTTTTGGGCTTTGC 1019
DB 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGCCCTTTTACCAATGTGGCTATCTCTGCTGTATGATGCTTTTATATGATGATATCAATC 1079
DB 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
QY 1080 TAAAGCAGGCTTTTCACTTTCTCGCAACTTCAAGGCTTTTCTGTGTAACAATATCTGAA 1139
DB 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi 683
QY 1140 CTTTATCCCGGTTGCCCGCAACGGTCCGCTCTCTGCAAGTGTCTGCTGACGCAACCCC 1199
DB 683 sLeuTyrProValAlaArg---ArgThrAlaLeuCysGlnValPheAlaAspAlaThrPr 702
QY 1200 CACTGATGGGGCTTGGCCATAGCCCATAGCGCTGCTGGAACCTTTCTGGCTCTCTCT 1259
DB 702 oThrGlyTyrGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 722
QY 1260 GCCGATCCATCTGCGGAACCTCTAGCAGCTTGTGCTGCGAGCGGCTCTGGAGCAAA 1319
DB 722 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy 742
QY 1320 ACTTATCGGAACCGCAACTCTGTGCTCTCTCGGAATATACACCTCTTCCATGGCT 1379
DB 742 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 762
QY 1380 GCTAGGTGTGCTGCGCAACTCGATCTGCGGGGAGCTCTCTGCTACGCTCCGCTCGGC 1439
DB 762 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrTyrPheValTyrValProSerAl 782
QY 1440 GCTGAATCCCGCGACGACCGCTCTCGGGCGGCTTTGGGGCTCTACCGTCCCTCTTCTCA 1499
DB 782 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuIleArgProLeuLeuHi 802
QY 1500 TCTGCGGTTCCGGCGACACCGCGCGCACCTCTCTTTACGCGTCTCCCGTATGTGCC 1559
DB 802 sLeuArgPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 822
QY 1560 TTCTCATCTGCGGACCGTGTGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1619
DB 822 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaIleTrpArgProPr 842
QY 1620 G 1620
DB 842 o 842

RESULT 13
US-08-591-502B-61
; Sequence 61, Application US/08591502B
; Patent No. 660727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
  
```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591.502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 832 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-08-591-502B-61

```

Alignment Scores:

```

Pred. No.: 7,84e-234 Length: 832
Score: 2652.00 Matches: 494
Percent Similarity: 95.38% Conservative: 22
Best Local Similarity: 91.31% Mismatches: 24
Query Match: 44.90% Indels: 2
DB: 2 Gaps: 0

```

US-10-761-006a-1 (1-3215) x US-08-591-502B-61 (1-832)

```

QY 1 CTCACGAATTCACCAACCTGCTAGATCCAGGGTCCAGGGGCTATATTTCTCTGC 60
DB 293 LeuHisAsnLeuProAsnSerAlaArgSerGlnSerGluArgProValPheProCys 312
QY 61 TGTGGCTCCAGTTCGGAAACAGTAAACCTGTTCCGACTACTGCTCTCCCATATCGTC 120
DB 313 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal 332
QY 121 AATCTTCTCAGAGCTGGGACCTCGACCGAATCATGGAGAACACATCAGGATTCCT 180
DB 333 AsnLeuLeuGluAspTrpGlyProCysAlaGluHisGlyGluHisIleArgIlePro 352
QY 181 AGGACCTCTCGTGTACAGGGGGTTCCTGTTGACAGAGATCCTCAATACC 240
DB 353 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 372
QY 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGGAGCACCCACGTGTTCC 299
DB 373 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnTyrArgVal-Ser 392
QY 300 CTGGCCAAAATTCGAGTCCCAACTCCAACTCAATCACTCAACCACTCTGTCCTCAATTT 359
DB 392 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 412
QY 360 GTCTGGCTATCCGTGATGTCGCGGGTTCATATTCCTTTCATCTCTCTCTCTCTCTCTCT 419
DB 412 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl 432
QY 420 TATGCTCATCTCTTGTGTTGGTCTCTCTGAGCTACCAAGGTATGTTGCGCGGTTGTCTCT 479
DB 432 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 452

```

```

QY 480 TACTTCCAGAAATCAACCAACGACGCGGGCCATGCAAGACCTGCACGACTCTCTGCTC 539
DB 452 rAsnSerArgIlePheAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 472
QY 540 AAGGAAACCTCTACGTTTCCCTCTGTTGTGTGTACAAAACCTTCGACGCGAAATCTGCACATT 599
DB 472 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrGlnThrPheGlyArgCysLeuHisLe 492
QY 600 GTATTCCCATCCCATCATCTCTGGGCTTTCGCAAGATTCTATGGGAGTGGGCTCTCAGTCC 659
DB 492 uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 512
QY 660 GTTTCCTCTGGCTCAGTTTACTAGTGCATTTGTTGTCAGTGGTTCGAGGGCTTCCCTCCA 719
DB 512 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 532
QY 720 CTGTTTGGCTTTTACCTCTATTACCAATTTCTTTTGTCTTTTGGGTATACATTTAAACCC 779
DB 532 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 552
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTTGGGTATACATTTAAACCC 839
DB 552 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 572
QY 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACCTTCACTGAGATATGTAATTTGGAAGTTG 899
DB 572 oAsnLysThrLysArgTrpGlyTyrSerLeuHisPheMetGlyTyrValIleGlyCysTy 592
QY 900 GGTACTTTTACCGCAGGAACATATTGTACTAAACCTCAAGCAATGTTTTCGAAAACCTGCC 959
DB 592 rGlySerLeuProGlnAspHisIleGlnLysIleLysGluCysPheArgLysLeuPr 612
QY 960 TGTAAATAGACCTATTGATTGGAAGATGATGTCAAAGAATTGTGGGTCTTTTGGGCTTTC 1019
DB 612 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 632
QY 1020 TGGCCCTTTTACAAATGTGGCTATCTCCCTTGATGCTTATATGATGATATATCAATC 1079
DB 632 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 652
QY 1080 TAAGCAGGCTTTTCACTTTCTCCCAACTTACAAAGGCTTTCTGTGTAAACAATATCTGAA 1139
DB 652 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 672
QY 1140 CCTTTACCCCGTTGCGCGCAACGGTCCGCTCTCTGCCAAGTGTGTTGCTGCAGCAACCCC 1199
DB 672 nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 692
QY 1200 CACTGGATGGGCTTGGCCATAGGCCATCAGGCATGGCTGGAACTTTCTGGCTCTCTCT 1259
DB 692 oThrGlyTrpGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheLeuAlaArgLe 712
QY 1260 GCGCATCCATCTCGCGAACTCTTAGCAGCTGTTGTTGCTCGCAGCGGTCTGAGGACAAA 1319
DB 712 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAs 732
QY 1320 ACTTATCGGAACCGCAACTCTGTTGTCTCTCTCGAAATACACCTCTTTTCCATGGCT 1379
DB 732 nIleLeuGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerTyrProTrpLe 752
QY 1380 GCTAGGCTGTCTGCCAACTGGATCTCGCGGGAGCTCTTTGTCTAGTCCCGTCCGCGC 1439
DB 752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 772
QY 1440 GCTGAATCCCGCGGACGACCCGCTCTCGGGCGGTGTTGGGGCTCTACCGTCCCTTCTCA 1499
DB 772 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuSerArgProLeuLeuAr 792
QY 1500 TCTGCGGTTCCGCGCGACACCGGGCGCACCTCTCTTTACGGGGTCTCCCGGTATGTGCC 1559
DB 792 gLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValPr 812

```

1560 TTCTCATCTCCGCGACCGTGTGACCTTCTGCTTCTACCTCTCAGCTGCGAGACACC 1619
812 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 832

1620 G 1620
832 o 832

RESULT 14
US-08-591-502B-62
Sequence 62, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:
APPLICANT: Chisari, Francis V.
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
Lymphocyte Responses to Hepatitis B Virus

NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08685
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Leuwer
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-0002300S
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 832 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 62:

US-08-591-502B-62
Alignment Scores:
Pred. No.: 2,78e-233 Length: 832
Score: 2646.00 Matches: 495
Percent Similarity: 95.01% Conservative: 19
Best Local Similarity: 91.50% Mismatches: 26
Query Match: 44.79% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x US-08-591-502B-62 (1-832)

1 CTCACAACTTCCACCAAGCTCTGCTAGATCCCAAGGCTGAGGGCCCTATATTTCTCTGC 60
233 LeuHisAsnLeuProAsnSerAlaArgSerGlnSerGluArgProValPheProCys 312

QY 61 TGGTGGCTCCAGTTCGGGAACAGTAAACCTGTTCGACTACTGCTCTCTCCCATATCGTC 120
DB 313 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerLeuIleVal 332
QY 121 AATCTTCTCAGACTGGGGACCCCTGCACCGAATGGAGACACACAACTCAGGATTCCT 180
DB 333 AsnLeuLeuGluAspTrpGlyProCysAlaGluHisGlyGluHisIleArgIlePro 352
QY 181 AGGACCCCTGCTGCTTACAGCGGGGTTTTCTCGTTGACAGAATCTCTCAATACC 240
DB 353 ArgThrProSerArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 372
QY 241 GCAGAGTCTAGACTCTG-GTGAGCTTCTCTCAATTTCTAGGGGGAGACCCACGTTCTC 299
DB 373 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnTyrArgVal-Se 392
QY 300 CTGGCCAAATTCGCAGTCCCAACCTCCAATCACCTCACCAACCTCTGTCTCTCCAAATTT 359
DB 392 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 412
QY 360 GTCCTGCTATGCTGATGTCTCGGGGTTTTATCATATTCTCTCTCTCTCTCTCTCTGC 419
DB 412 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl 432
QY 420 TATGCTCATCTCTTGTGTTGTTCTTCTGACTACCAAGGTATGTGCTGCTCTCTCTCTC 479
DB 432 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 452
QY 480 TACTTTCAGGAACATCAACACGAGCAGCGGGCCATGCAAGACTGCAACACTGCTCTCTCTC 539
DB 452 rAsnSerArgIleLeuAsnHisGlnHisGlyThrMetProAsnLeuHisAspSerCysSe 472
QY 540 ARGMAACTCTAGCTTTCCTCTGTTGCTGTACAAACCTTCGGACGGAACCTGCACTT 599
DB 472 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrGlnThrPheGlyArgLysLeuHisLe 492
QY 600 GTATTCCTCATCTCTCTGCTTTCGCAAGATTCCTATGGAGTGGGCTCTCAGTCTC 659
DB 492 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 512
QY 660 GTTCTCTCTGCTCAGTTTACTAGTGCATTTGTTTCTAGTGGTTCGTAGGCTTTCCCTCA 719
DB 512 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 532
QY 720 CTGTTTGGCTTTCAGTTTATATGATGATGTTGTTGCGGCGAAGTCTGTACAACTCT 779
DB 532 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 552
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTTCTTGTCTTGTGGGTATACATTTAAACCC 839
DB 552 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 572
QY 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACCTTCACTTATGGGATATGTAATGGAAGTTG 899
DB 572 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlyCysTy 592
QY 900 GGGTACTTTTACCGCAGGAACATATTGTTACTTAAACTCAAGCAATGTTTTCGAAACTGCC 959
DB 592 rGlySerLeuProGlnGluHisIleIleGlnLysIleLysGluCysPheArgLysLeuPr 612
QY 960 TGTAAATAGACCTATTGATTGGAAAGTATGTCAAGAAATTTGTTGGTCTTTTGGGCTTTGC 1019
DB 612 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 632
QY 1020 TGCCCCCTTTTACAAATGTGGCTATCTGCTGCTTATGCTGCTTATGATGCTATGATCAATC 1079
DB 632 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 652
QY 1080 TAAGCAGCTTTCAGTTTCTCGCCCAACTTACAAGSCCTTTCTGTGTAAACAATATCTCAA 1139
DB 652 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 672
QY 1140 CTTTACCCTTCCCGGCAACGGTCCGGTCTCTGCGCAAGTGTGTTGTGACGCAACCCC 1199

```
Db      672 nLeuTyPProValAlaGlyGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 692
QY      1200 CACTGGATGGGCTTGCCATAGCCATCAGCCATGCTGGACCTTTCTGGCTCTCT 1259
Db      692 oThrGlyTrpGlyLeuAlaMetGlyHisGlnArgMetArgGlyThrPheSerAlaProLe 712
QY      1260 GCCGATCCATACGCGGAACTCCTAGCAGCTTGTGTGCTCGCAGCGGCTCGAGCAAA 1319
Db      712 uProIleHisThrAlaGlnLeuLeuAlaLaCysPheAlaArgSerArgSerGlyAlaAs 732
QY      1320 ACTTATCGGAACCGAACAACCTCTGTGTCTCTCTCGGAATACACCTCTCTTCCATGGCT 1379
Db      732 nIleIleGlyThrAspAsnSerValValLeuSerArgLysTyThrSerPheProTrie 752
QY      1380 GCTAGGCTGTGTCGCAACTGATCTCGCGGAGCTCTTTGTCTACGTCGCCGTCGGC 1439
Db      752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyValProSerAl 772
QY      1440 GCTGAATCCCGCGAGCACCGCTCTCGGGCCGTTTGGGGCTCTACCGTCCCTTCTTCA 1499
Db      772 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuSerArgProLeuLeuAr 792
QY      1500 TCTGCGCTTCCGCGCCAGCACCGGCGCACTCTCTTTAGCGGTCTCCCGGTATGTGC 1559
Db      792 gLeuProPheArgProThrThrGlyArgThrSerLeuTyAlaAspSerProSerValPr 812
QY      1560 TTCTATCTGCCGAGCGGTGTGCACTTCGCTTCACCTCTGACGTCGCATGGAGACCACC 1619
Db      812 oSerHisLeuProAspLeuValHisPheAlaSerProLeuHisValAlaAlaTrpArgProPr 832
QY      1620 G 1620
Db      832 o 832

RESULT 15
US-08-591-502B-56
; Sequence 56, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
```

```
;
; TELCOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 843 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 56:
US-08-591-502B-56

Alignment Scores:      3.46e-233      Length:      843
Pred. No.:             2645.00        Matches:     493
Percent Similarity:    95.19%         Conservative: 22
Best Local Similarity: 91.13%         Mismatches:  25
Query Match:           44.78%         Indels:       2
DB:                    2              Gaps:         0

US-10-761-006A-1 (1-3215) x US-08-591-502B-56 (1-843)

QY      1 CTCCACAACATTCCACCAAGCTCTGTAGATCCAGGGTGAGGGGCTATATTTTCTCTGC 60
Db      304 LeuHisHisValProAsnSerSerArgSerGlnSerGlnGlySerValLeuSerCys 323
QY      61 TGGTGGCTCCAGTTCGGAACAGTAACCCCTGTTCGAGCTACTGCTCTCCCATATCGTC 120
Db      324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerGluHisCysLeuPheHisIleVal 343
QY      121 AATCTTCTCAGGACTGGGGACCCCTGCACCGAACATCGGAGAACACAAATCAGGATTCCT 180
Db      344 AsnLeuIleGluAspTrpGlyProCysAlaGluHisGlyGluHisArgIleArgThrPro 363
QY      181 AGGACCCCTGCTCGTGTACAGCGGGGTTTTCTCGTTGACAAAGAAATCCTCAACAATACC 240
Db      364 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 393
QY      241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGGAGACCCACGCTGTTTC 299
Db      384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnThrArgVal-Se 403
QY      300 CTGGCCAAAATTGCGAGTCCCAACCTCCAATCACTCAACAACCTCTCTCTCCCAATTT 359
Db      403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAspLe 423
QY      360 GTCTTGGCTATCGCTGGATGTGTCTGGCGGTTTTATCATATTTCTCTTTCATCTCTGCTGC 419
Db      423 userTrpLeuSerLeuAspValSerAlaAlaPheTyHisLeuProLeuHisProAlaAl 443
QY      420 TATGCTCATCTTCTTGTGTTGTTCTGTGAGTACCAAGGTATGTTGCCGTTGTGCTTC 479
Db      443 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyTrValAlaArgLeuSerSe 463
QY      480 TACTTCCAGGAACATCAACCACGACGCGGGCCATGCAAGACCTGCACGACTCTCTGCTC 539
Db      463 rAsnSerArgIleIleAsnHisGlnHisArgThrMetGlnAsnLeuHisAspSerCysSe 483
QY      540 AAGGAAACTCTACGTTTCCCTCTGTGTGTGTACAAAAACCTTCGGACGCGAAATCGCACTT 599
Db      483 rArgAsnLeuTyTrValSerLeuLeuLeuTyTrLysThrTyGlyArgLysLeuHisLe 503
QY      600 GTATTCCCATCCCATCATCTCTGGGCTTTTCGCAAGATTCCTATGGAGTGGGGCTCAGTCC 659
Db      503 uTySerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
QY      660 GTTTCCTCGGCTCAGTTTACTAGTGCACATTTGTTGTCAGTGGTTCGTAGGGCTTCCCCCA 719
Db      523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
QY      720 CTGTTTGGCTTTTCAGTTATATGATGATGTGTATTTGGGGGCGAAGTCTGTACAAATCT 779
```


Db 543 sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTGCTCTTTGGGTATACATTTAAACCC 839
Db 563 uGluSerLeuTyrAlaAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
QY 840 TAATAAACCAACGGTTGGGCTACTCCCTTAACCTCATGGGATATGTAATTTGGAAGTTG 899
Db 583 oGlnLysThrLysArgTrpGlyTyrSerLeuAsnPhenMetGlyTyrValIleGlySerTr 603
QY 900 GGTACTCTTTACCGCAGGAACATATTGTACTAAACTCAAGCAATGTTTTCGAAAACTGCC 959
Db 603 pGlyThrLeuProGlnGluHisIleValLeuLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAAATAGACCTATTGATTGGAAGATATGTCAAGAAATTGTGGGCTCTTTGGGCTTGC 1019
Db 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGCCCTTTTACACAATGGCTATCCTGCTTGCCTTATATGCTTATATGCTATATACAAATC 1079
Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnAl 663
QY 1080 TAAGCAGGCTTTCACTTTCTCGCAACTTACAAGGCTTTCTGTGTAAACAATATCTGAA 1139
Db 663 aLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuThrLysGlnTyrLeuAs 683
QY 1140 CCTTTACCCGTTGCGCGCAACGGTTCGGTCTCTGCAAGTGTTCCTGACGCAACCCC 1199
Db 683 nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 703
QY 1200 CACTGGATGGGCTTGGCCATAGCCATCAGCGATGCTGGAACCTTTCTGGCTCTCT 1259
Db 703 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValSerProLe 723
QY 1260 GCCGATCCATACCTCGGAACCTCTAGCAGCTTTGTTTTCTCGCAGCGGCTCTGGAGCAA 1319
Db 723 uProIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLys 743
QY 1320 ACTTATCGGAACCGCAACTCTGTTGCTCTCTCGGAAATACACCTCTCTTCCATGGCT 1379
Db 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
QY 1380 GCTAGGTGCTGCTCCAACTGGATCCTCGCGGAGCTCTTTGTACTACGTCCGTCGGC 1439
Db 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
QY 1440 GCTGAATCCCGGACGACCGCTCTCGGGGCGGTTTGGGGCTCTACCGTCCCTTCTTCA 1499
Db 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuAr 803
QY 1500 TCTGCGTTTCCGGCGGACGACGCGGCGGACCTCTCTTTACGGGCTCTCCCGTATGTGCC 1559
Db 803 gLeuProTyrArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValPr 823
QY 1560 TTCTCATCTGCGGACCGCTGCTGCTCTGCTTCCCTCTGACGCTCGCATGGAGACCACC 1619
Db 823 oSerArgLeuProAspArgValHisPheAlaSerProLeuHisValAlaIleTrpArgProPr 843
QY 1620 G 1620
Db 843 O 843

